

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
19 October 2006 (19.10.2006)

PCT

(10) International Publication Number
WO 2006/110581 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01) *G01N 33/574* (2006.01)

(US). **MOLER, Edward** [US/US]; Walnut Creek, CA
(US). **ROWE, Michael** [US/US]; Oakland, CA (US).

(21) International Application Number:
PCT/US2006/013153

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(22) International Filing Date: 7 April 2006 (07.04.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/669,861 7 April 2005 (07.04.2005) US

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicants (*for all designated States except US*): **CH-IRON CORPORATION** [US/US]; P.O. Box 8097, Emeryville, CA 94602-8097 (US). **SAGRES DISCOVERY, INC.** [US/US]; P.O. Box 8097, Emeryville, CA 94602-8097 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **LAI, Albert** [CA/US]; Davis, CA (US). **FANIDI, Abdallah** [US/US]; Fair Oaks, CA (US). **BOOHER, Robert** [US/US]; Davis, CA (US). **TSE, Christin** [US/US]; 1321 Saint William Drive, Libertyville, IL 60048 (US). **XU, Xie** [CN/US]; 6500 Paseo Padre Parkway, Mailstop B2f2se, Fremont, CA 94555 (US). **YU, Guoying** [US/US]; Kensington, CA

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CANCER-RELATED GENES

(57) Abstract: This invention is in the field of cancer-associated genes. Specifically it relates to methods for detecting cancer or the likelihood of developing cancer based on the presence or absence of expression of certain genes or proteins encoded by those genes. The invention also provides methods and molecules for upregulating or downregulating these cancer-associated genes.



WO 2006/110581 A2

CANCER-RELATED GENES

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims priority of U.S. Serial No. 60/669,861, filed April 7, 2005, which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

[0002] This invention is in the field of cancer-associated genes. Specifically it relates to methods for detecting cancer or the likelihood of developing cancer based on the presence of differential expression of certain genes or their gene products. The invention also provides methods and molecules for detecting, diagnosing and treating cancer by modulating these cancer-associated genes.

BACKGROUND OF THE INVENTION

[0003] Oncogenes are genes that can cause cancer. Carcinogenesis can occur by a wide variety of mechanisms, including infection of cells by viruses containing oncogenes, activation of protooncogenes (normal genes that have the potential to become an oncogene) in the host genome, and mutations of protooncogenes and tumour suppressor genes. Carcinogenesis is fundamentally driven by somatic cell evolution (*i.e.* mutation and natural selection of variants with progressive loss of growth control). The genes that serve as targets for these somatic mutations are classified as either protooncogenes or tumour suppressor genes, depending on whether their mutant phenotypes are dominant or recessive, respectively.

[0004] There are a number of viruses known to be involved in human as well as animal cancer. Of particular interest here are viruses that do not contain oncogenes themselves; these are slow-transforming retroviruses. Such viruses induce tumours by integrating into the host genome and affecting neighboring protooncogenes in a variety of ways. Provirus insertion mutation is a normal consequence of the retroviral life cycle. In infected cells, a DNA copy of the retrovirus genome (called a provirus) is integrated into the host genome. A newly integrated provirus can affect gene expression in *cis* at or near the integration site by one of two mechanisms. Type I insertion mutations up-regulate transcription of proximal genes as a consequence of regulatory sequences (enhancers and/or promoters) within the proviral long terminal repeats (LTRs). Type II insertion mutations located within the intron or exon of a gene can up-regulate transcription of said gene as a consequence of regulatory sequences (enhancers and/or promoters) within the

proviral long terminal repeats (LTRs). Additionally, type II insertion mutations can cause truncation of coding regions due to either integration directly within an open reading frame or integration within an intron flanked on both sides by coding sequences, which could lead to a truncated or an unstable transcript/protein product. The analysis of sequences at or near the insertion sites has led to the identification of a number of new protooncogenes.

[0005] With respect to lymphoma and leukemia, retroviruses such as AKV murine leukemia virus (MLV) or SL3-3 MLV, are potent inducers of tumours when inoculated into susceptible newborn mice, or when carried in the germline. A number of sequences have been identified as relevant in the induction of lymphoma and leukemia by analyzing the insertion sites; see Sorensen *et al.*, J. Virology 74:2161 (2000); Hansen *et al.*, Genome Res. 10(2):237-43 (2000); Sorensen *et al.*, J. Virology 70:4063 (1996); Sorensen *et al.*, J. Virology 67:7118 (1993); Joosten *et al.*, Virology 268:308 (2000); and Li *et al.*, Nature Genetics 23:348 (1999); all of which are expressly incorporated by reference herein. With respect to cancers, especially breast cancer, prostate cancer and cancers with epithelial origin, the mammalian retrovirus, mouse mammary tumour virus (MMTV) is a potent inducer of tumours when inoculated into susceptible newborn mice, or when carried in the germ line. *Mammary Tumours in the Mouse*, edited by J. Hilgers and M. Sluyser; Elsevier/North-Holland Biomedical Press; New York, N.Y.

[0006] The pattern of gene expression in a particular living cell is characteristic of its current state. Nearly all differences in the state or type of a cell are reflected in the differences in RNA levels of one or more genes. Comparing expression patterns of uncharacterized genes may provide clues to their function. High throughput analysis of expression of hundreds or thousands of genes can help in (a) identification of complex genetic diseases, (b) analysis of differential gene expression over time, between tissues and disease states, and (c) drug discovery and toxicology studies. Increase or decrease in the levels of expression of certain genes correlate with cancer biology. For example, oncogenes are positive regulators of tumourigenesis, while tumour suppressor genes are negative regulators of tumourigenesis. (Marshall, Cell, 64: 313-326 (1991); Weinberg, Science, 254: 1138-1146 (1991)).

[0007] Immunotherapy, or the use of antibodies for therapeutic purposes has been used in recent years to treat cancer. Passive immunotherapy involves the use of monoclonal antibodies in cancer treatments. See for example, Cancer: Principles and Practice of Oncology, 6th Edition (2001) Chapt. 20 pp. 495-508. Inherent therapeutic biological activity of these antibodies include direct inhibition of tumour cell growth or survival, and the ability to recruit the natural cell killing activity of the body's immune system. These agents are administered alone or in

conjunction with radiation or chemotherapeutic agents. Rituxan® and Herceptin®, approved for treatment of lymphoma and breast cancer, respectively, are two examples of such therapeutics. Alternatively, antibodies are used to make antibody conjugates where the antibody is linked to a toxic agent and directs that agent to the tumour by specifically binding to the tumour. Mylotarg® is an example of an approved antibody conjugate used for the treatment of leukemia. However, these antibodies target the tumour itself rather than the cause.

[0008] An additional approach for anti-cancer therapy is to target the protooncogenes that can cause cancer. Genes identified as causing cancer can be monitored to detect the onset of cancer and can then be targeted to treat cancer.

SUMMARY OF THE INVENTION

[0009] In some aspects, the present invention provides methods for treating cancer in a patient comprising modulating the level of an expression product of a cancer-associated gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA. In some embodiments the cancer is lymphoma, melanoma, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, bladder cancer, stomach cancer or skin cancer.

[0010] In some aspects, the present invention provides methods of treating a cancer in a patient characterized by overexpression of the gene relative to a control. In some embodiments the method comprises modulating gene expression in the patient.

[0011] In some aspects, the present invention provides methods for diagnosing cancer comprising detecting evidence of differential expression in a patient sample of a cancer-associated gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP,

IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA. In some embodiments evidence of differential expression of the gene is diagnostic of cancer.

[0012] In some aspects, the present invention provides methods for detecting a cancerous cell in a patient sample comprising detecting evidence of an expression product of a cancer-associated gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA. In some embodiments evidence of expression of the gene in the sample indicates that a cell in the sample is cancerous.

[0013] In some aspects, the present invention provides methods for assessing the progression of cancer in a patient comprising comparing the level of an expression product of a cancer-associated gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA in a biological sample at a first time point to a level of the same expression product at a second time point. In some embodiments a change in the level of the expression product at the second time point relative to the first time point is indicative of the progression of the cancer.

[0014] In some aspects, the present invention provides methods of diagnosing cancer comprising:

(a) measuring a level of mRNA of a cancer-associated gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, in a first sample, said first sample comprising a first tissue type of a first individual; and

(b) comparing the level of mRNA in (a) to:

(1) a level of the mRNA in a second sample, said second sample comprising a normal tissue type of said first individual, or

(2) a level of the mRNA in a third sample, said third sample comprising a normal tissue type from an unaffected individual. In some embodiments at least a two fold difference between the level of mRNA in (a) and the level of the mRNA in the second sample or the third sample indicates that the first individual has or is predisposed to cancer.

[0015] In some aspects, the present invention provides of screening for anti-cancer activity comprising:

(a) contacting a cell that expresses a cancer-associated gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, with a candidate anti-cancer agent; and

(b) detecting at least a two fold difference between the level of the gene's expression in the cell in the presence and in the absence of the candidate anti-cancer agent. In some embodiments at least a two fold difference between the level of the gene's expression in the cell in the presence and in the absence of the candidate anti-cancer agent indicates that the candidate anti-cancer agent has anti-cancer activity.

[0016] In some aspects, the present invention provides methods for identifying a patient as susceptible to treatment with an antibody that binds to an expression product of a cancer-associated gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, comprising measuring the level of the expression product of the gene in a biological sample from that patient.

[0017] In some aspects, the present invention provides kit for the diagnosis or detection of cancer in a mammal. In some embodiments the kit comprises an antibody or fragment thereof, or an immunoconjugate or fragment thereof, according to any one of the proceeding embodiments. In some embodiments the antibody or fragment specifically binds a cancer-associated tumor cell antigen; one or more reagents for detecting a binding reaction between said antibody and said tumor cell antigen. In some embodiments the kits comprise instructions for using the kit.

[0018] In some aspects, the present invention provides kits for diagnosing cancer comprising a nucleic acid probe that hybridises under stringent conditions to a cancer-associated gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, and primers for amplifying the gene. In some embodiments the kits comprise instructions for using the kit.

[0019] In some aspects, the present invention provides compositions comprising one or more antibodies or oligonucleotides specific for an expression product of a cancer-associated gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2,

CDH23, IFNAR2, RAMP1, MET, FGFR1, PLA2, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA.

[0020] These and other aspects of the present invention will be elucidated in the following detailed description of the invention.

DETAILED DESCRIPTION

[0021] Protooncogenes have been identified in humans using a process known as "provirus tagging", in which slow-transforming retroviruses that act by an insertion mutation mechanism are used to isolate protooncogenes using mouse models. In some models, uninfected animals have low cancer rates, and infected animals have high cancer rates. It is known that many of the retroviruses involved do not carry transduced host protooncogenes or pathogenic transacting viral genes, and thus the cancer incidence must therefore be a direct consequence of proviral integration effects into host protooncogenes. Since proviral integration is random, rare integrants will "activate" host protooncogenes that provide a selective growth advantage, and these rare events result in new proviruses at clonal stoichiometries in tumors. In contrast to mutations caused by chemicals, radiation, or spontaneous errors, protooncogene insertion mutations can be easily located by virtue of the fact that a convenient-sized genetic marker of known sequence (the provirus) is present at the site of mutation. Host sequences that flank clonally integrated proviruses can be cloned using a variety of strategies. Once these sequences are in hand, the tagged protooncogenes can be subsequently identified. The presence of provirus at the same locus in two or more independent tumors is *prima facie* evidence that a protooncogene is present at or very near the provirus integration sites (Kim et al, Journal of Virology, 2003, 77:2056-2062; Mikkers, H and Berns, A, Advances in Cancer Research, 2003, 88:53-99; Keoko et al. Nucleic Acids Research, 2004, 32:D523-D527). This is because the genome is too large for random integrations to result in observable clustering. Any clustering that is detected is unequivocal evidence for biological selection (i.e. the tumor phenotype). Moreover, the pattern of proviral integrants (including orientations) provides compelling positional information that makes localization of the target gene at each cluster relatively simple. The three mammalian

retroviruses that are known to cause cancer by an insertion mutation mechanism are FeLV (leukemia/lymphoma in cats), MLV (leukemia/lymphoma in mice and rats), and MMTV (mammary cancer in mice). Once protooncogenes have been identified in mouse models, the human orthologs can be annotated as protooncogenes and further investigations carried out.

[0022] Thus, the use of oncogenic retroviruses, whose sequences insert into the genome of the host organism resulting in cancer, allows the identification of host genes involved in cancer. These sequences may then be used in a number of different ways, including diagnosis, prognosis, screening for modulators (including both agonists and antagonists), antibody generation (for immunotherapy and imaging), etc. However, as will be appreciated by those in the art, oncogenes that are identified in one type of cancer such as those identified in the present invention, have a strong likelihood of being involved in other types of cancers as well.

[0023] The invention therefore provides methods for detecting cancerous cells in a biological sample comprising determining the sequence or expression level of one or more (i.e. 1, 2, 3, 4, 5 or more) cancer-associated genes.

[0024] As used herein, the term "cancer-associated genes" refers to genes selected from WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA.

[0025] These genes have been identified and validated as proto-oncogenes using the methods described herein.

[0026] In some embodiments the methods include measuring the level of expression of one or more (i.e. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) expression products of the cancer-associated gene, wherein a level of expression that is different to a control level is indicative of disease.

[0027] In some embodiments the expression product is a protein, although alternatively mRNA expression products may be detected. If a protein is used, the protein is preferably detected by an antibody which preferably binds specifically to that protein. The term "binds specifically" means that the antibodies have substantially greater affinity for their target polypeptide than their affinity for other related polypeptides. As used herein, the term "antibody" refers to intact

molecules as well as to fragments thereof, such as Fab, F(ab')₂ and Fv, which are capable of binding to the antigenic determinant in question. By “substantially greater affinity” we mean that there is a measurable increase in the affinity for the target polypeptide of the invention as compared with the affinity for other related polypeptide. In some embodiments, the affinity is at least 1.5-fold, 2-fold, 5-fold 10-fold, 100-fold, 10³-fold, 10⁴-fold, 10⁵-fold, 10⁶-fold or greater for the target polypeptide.

[0028] In some embodiments, the antibodies bind with high affinity, with a dissociation constant of 10⁻⁴M or less, 10⁻⁷M or less, 10⁻⁹M or less; or subnanomolar affinity (0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1 nM or even less).

[0029] Where mRNA expression product is used, in some embodiments it is detected by contacting a tissue sample with a probe under conditions that allow the formation of a hybrid complex between the mRNA and the probe; and detecting the formation of a complex. In some embodiments stringent hybridization conditions are used.

[0030] Cancer associated genes themselves may be detected by contacting a biological sample with a probe under conditions that allow the formation of a hybrid complex between a nucleic acid expression product encoding a cancer-associated gene and the probe; and detecting the formation of a complex between the probe and the nucleic acid from the biological sample. In some embodiments, the absence of the formation of a complex is indicative of a mutation in the sequence of the cancer-associated gene.

[0031] Methods include comparing the amount of complex formed with that formed when a control tissue is used, wherein a difference in the amount of complex formed between the control and the sample indicates the presence of cancer. In some embodiments the difference between the amount of complex formed by the test tissue compared to the normal tissue is an increase or decrease. In some embodiments a two-fold increase or decrease in the amount of complex formed is indicative of disease. In some embodiments, a 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold or even 100-fold increase or decrease in the amount of complex formed is indicative of disease.

[0032] In some embodiments the biological sample used in the methods of the invention is a tissue sample. Any tissue sample may be used. In some embodiments, however, the tissue is selected from breast tissue, colon tissue, kidney tissue, liver tissue, lung tissue, lymphoid tissue, ovary tissue, pancreas tissue, prostate tissue, uterine tissue, cervix tissue or skin tissue.

[0033] The invention also provides methods for assessing the progression of cancer in a patient comprising comparing the expression of one or more (i.e. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more)

expression products of the cancer-associated genes referred to above in a biological sample at a first time point to the expression of the same expression product at a second time point, wherein an increase or decrease in expression, or in the rate of increase or decrease of expression, at the second time point relative to the first time point is indicative of the progression of the cancer.

[0034] The invention also provides kits useful for diagnosing cancer comprising an antibody that binds to a polypeptide expression product of a cancer-associated gene; and a reagent useful for the detection of a binding reaction between said antibody and said polypeptide. In some embodiments, the antibody binds specifically to the polypeptide product of the cancer-associated gene.

[0035] Furthermore, the invention provides a kit for diagnosing cancer comprising a nucleic acid probe that hybridises under stringent conditions to a cancer-associated gene; primers useful for amplifying the cancer-associated gene; and, optionally, instructions for using the probe and primers for facilitating the diagnosis of disease.

[0036] The invention further provides antibodies, nucleic acids, or proteins suitable for use in modulating the expression of an expression product of one or more (i.e. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) of the cancer-associated genes listed above, for use in treating cancer.

[0037] Accordingly, the invention provides methods for treating cancer in a patient, comprising modulating the level of one or more (i.e. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) expression products of any one of the cancer-associated genes listed above. In some embodiments the methods comprise administering to the patient a therapeutically-effective amount of an antibody, a nucleic acid, or a polypeptide that modulates the level of said expression product.

[0038] The invention therefore also provides the use of an antibody, a nucleic acid, or a polypeptide that modulates the level of an expression product of one or more (i.e. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) cancer-associated genes, in the manufacture of a medicament for the treatment, detection or diagnosis of cancer. In some embodiments the level of expression is modulated by action on the gene, mRNA or the encoded protein. In some embodiments the expression is upregulated or downregulated. For example, the change in regulation may be 2-fold, 3-fold, 5-fold, 10-fold, 20-fold, 50-fold, or even 100 fold or more.

[0039] Antibodies suitable for use in accordance with the present invention may be specific for cancer-associated proteins as these are expressed on or within cancerous cells. For example, glycosylation patterns in cancer-associated proteins as expressed on cancerous cells may be

different to the patterns of glycosylation in these same proteins as these are expressed on non-cancerous cells. In some embodiments antibodies according to the invention are specific for cancer-associated proteins as expressed on cancerous cells only. This is of particular value for therapeutic antibodies. Anti-target antibodies may also bind to splice variants, deletion, addition and/or substitution mutants of the target.

[0040] Antibodies suitable for therapeutic use in accordance with the present invention elicit antibody-dependent cellular cytotoxicity (ADCC). ADCC refers to the cell-mediated reaction wherein non-specific cytotoxic cells that express Fc receptors recognize bound antibody on a target cell and subsequently cause lysis of the target cell (Raghavan et al., 1996, *Annu Rev Cell Dev Biol* 12:181-220; Ghetie et al., 2000, *Annu Rev Immunol* 18:739-766; Ravetch et al., 2001, *Annu Rev Immunol* 19:275-290). Antibodies suitable for therapeutic use in accordance with the present invention may elicit antibody-dependent cell-mediated phagocytosis (ADCP). ADCP is the cell-mediated reaction wherein nonspecific cytotoxic cells that express Fc receptors recognize bound antibody on a target cell and subsequently cause phagocytosis. These processes are mediated by natural killer (NK) cells, which possess receptors on their surface for the Fc portion of IgG antibodies. When IgG is made against epitopes on "foreign" membrane-bound cells, including cancer cells, the Fab portions of the antibodies react with the cancerous cell. The NK cells then bind to the Fc portion of the antibody.

[0041] In embodiments where it is desirable to modify the antibody of the invention with respect to effector function, e.g. so as to enhance antigen-dependent cell-mediated cytotoxicity (ADCC) and/or complement dependent cytotoxicity (CDC) of the antibody, one or more amino acid substitutions can be introduced into an Fc region of the antibody. Alternatively or additionally, cysteine residue(s) may be introduced in the Fc region, thereby allowing interchain disulfide bond formation in this region (For review: Weiner and Carter (2005) *Nature Biotechnology* 23(5): 556-557). The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., *J. Exp Med.* 176:1191-1195 (1992) and Shopes, B. J. *Immunol.* 148:2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff et al. *Cancer Research* 53:2560-2565 (1993). Alternatively, an antibody can be engineered which has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al. *Anti-Cancer Drug Design* 3:219-230 (1989). Antibodies can be produced with modified glycosylation within the Fc region. For example, lowering the fucose content in

the carbohydrate chains may improve the antibody's intrinsic ADCC activity (see for example BioWa's Potillegent™ ADCC Enhancing Technology, described in WO0061739). Alternately, antibodies can be produced in cell lines that add bisected non-fucosylated oligosaccharide chains (see US 6,602,684). Both these technologies produce antibodies with an increased affinity for the FcγRIIIa receptor on effector cells which results in increased ADCC efficiency. The Fc region can also be engineered to alter the serum half life of the antibodies of the invention. Abdegs are engineered IgGs with an increased affinity for the FcRn salvage receptor, and so have shorter half life than conventional IgGs (see Vaccaro et al, (2005) Nature Biotechnology 23(10): 1283- 1288). To increase serum half life, specific mutations can be introduced into the Fc region that appear to decrease the affinity with FcRn (see Hinton et al, (2004) J Biol Chem 297(8): 6213-6216). Antibodies of the invention can also be modified to use other mechanisms to alter serum half life, such as including a serum albumin binding domain (dAb) (see WO05035572 for example). Engineered Fc domains (see for example XmAB™, WO05077981) may also be incorporated into the antibodies of the invention to lead to improved ADCC activity, altered serum half life or increased antibody protein stability.

[0042] In some embodiments, antibodies for therapeutic use in accordance with the invention are effective to elicit ADCC, and modulates the survival of cancerous cells by binding to target and having ADCC activity. Antibodies can be engineered to heighten ADCC activity (see, for example, US 20050054832A1, Xencor Inc. and the documents cited therein).

[0043] In some embodiments the nucleic acid type used in such methods is an antisense construct, a ribozyme or RNAi, including, for example, siRNA.

[0044] The cancer may be treated by the inhibition of tumour growth or the reduction of tumour volume or, alternatively, by reducing the invasiveness of a cancer cell. In some embodiments, the methods of treatment described above are used in conjunction with one or more of surgery, hormone ablation therapy, radiotherapy or chemotherapy. For example, if a patient is already receiving chemotherapy, a compound of the invention that modulates the level of an expression product as listed above may also be administered. The chemotherapeutic, hormonal and/or radiotherapeutic agent and compound according to the invention may be administered simultaneously, separately or sequentially.

[0045] In some embodiments the cancer being detected or treated according to one of the methods described above is selected from breast cancer, colon cancer, kidney cancer, liver

cancer, lung cancer, lymphoid cancer, ovary cancer, pancreas cancer, prostate cancer, uterine cancer, cervix cancer or skin cancer.

[0046] The invention provides methods for diagnosing cancer comprising detecting evidence of differential expression in a patient sample of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA. Evidence of differential expression of the gene is diagnostic of cancer. In some embodiments the cancer is lymphoma, leukemia, melanoma, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, bladder cancer, stomach cancer or skin cancer. In some embodiments, evidence of differential expression of the gene is detected by measuring the level of an expression product of the gene. In some embodiments the expression product is a protein or mRNA. In some embodiments the level of expression of protein is measured using an antibody which binds specifically to the protein. In some embodiments the antibody is linked to an imaging agent. In some embodiments the level of expression product of the gene in the patient sample is compared to a control. In some embodiments the control is a known normal tissue of the same tissue type as in the patient sample. In some embodiments the level of the expression product in the sample is increased relative to the control.

[0047] The invention also provides methods for detecting a cancerous cell in a patient sample comprising detecting evidence of an expression product of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA. Evidence of expression of the gene in the sample indicates that a cell in the sample is cancerous.

In some embodiments the cell is a breast cell, colon cell, kidney cell, liver cell, lung cell, lymphatic cell, ovary cell, pancreas cell, prostate cell, uterine cell, cervical cell, bladder cell, stomach cell or skin cell. In some embodiments evidence of the expression product is detected using an antibody linked to an imaging agent.

[0048] The invention provides methods for assessing the progression of cancer in a patient comprising comparing the level of an expression product of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFB3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, in a biological sample at a first time point to a level of the same expression product at a second time point. A change in the level of the expression product at the second time point relative to the first time point is indicative of the progression of the cancer. In some embodiments the cancer is lymphoma, leukemia, melanoma, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, bladder cancer, stomach cancer or skin cancer.

[0049] The invention also provides methods of diagnosing cancer comprising (a) measuring a level of mRNA of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFB3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, in a first sample wherein the first sample comprises a first tissue type of a first individual; and (b) comparing the level of mRNA in (a) to a control. Detection of at least a two fold difference between the level of mRNA in (a) and the level of the mRNA in the second sample or the third sample indicates that the first individual has or is predisposed to cancer. In some embodiments the control sample comprises a normal tissue type of the first individual. In some embodiments the control sample comprises a normal

tissue type from an unaffected individual. In some embodiments, at least a three fold difference between the level of mRNA in the first sample and the control indicates that the first individual has or is predisposed to cancer.

[0050] The invention provides methods of screening for anti-cancer activity comprising (a) contacting a cell that expresses a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTBR, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, with a candidate anti-cancer agent; and (b) detecting at least a two fold difference between the level of gene expression in the cell in the presence and in the absence of the candidate anti-cancer agent. At least a two fold difference between the level of gene expression in the cell in the presence compared to the level level of gene expression in the cell in the absence of the candidate anti-cancer agent indicates that the candidate anti-cancer agent has anti-cancer activity. In some embodiments at least a three fold difference between the level of gene expression in the cell in the presence and in the absence of the candidate anti-cancer agent indicates that the candidate anti-cancer agent has anti-cancer activity. In some embodiments the candidate anti-cancer agent is an antibody, small organic compound, small inorganic compound, or polynucleotide. In some embodiments the candidate anti-cancer agent is a monoclonal antibody. In some embodiments the candidate anti-cancer agent is a human or humanized antibody. In some embodiments the polynucleotide is an antisense oligonucleotide. In some embodiments the polynucleotide is an oligonucleotide having a sequence selected from the group consisting of SEQ ID NOS:1-144.

[0051] The invention also provides kits for the diagnosis or detection of cancer in a mammal. In some embodiments the kit comprises an antibody or fragment thereof, or an immunoconjugate or fragment thereof. In some embodiments the antibody or fragment is capable of specifically binding a tumor cell antigen wherein said tumor cell antigen is selected from the group of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2,

UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA. The kits further comprise one or more reagents for detecting a binding reaction between the antibody and the tumor cell antigen. In some embodiments the kit comprises instructions for using the kit.

[0052] The invention also provides kits for diagnosing cancer. In some embodiments the kits comprise a nucleic acid probe that hybridises under stringent conditions to a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAUI, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA. The kits also comprise primers for amplifying the cancer-associated gene. In some embodiments the kits comprise instructions for using the kit.

[0053] The invention provides methods for treating cancer in a patient. In some embodiments the methods comprises modulating the level of an expression product of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAUI, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA. In some embodiments the methods comprise administering to the patient an antibody, a nucleic acid, or a polypeptide that modulates the level of the expression product. In some embodiments the level of the expression product is upregulated or downregulated by at least a 2-fold change. In some embodiments the cancer is treated by the inhibition of tumour growth or the reduction of tumour volume. In some embodiments the cancer is treated by reducing the invasiveness of a cancer cell. In some embodiments the expression product is a protein or mRNA. In some embodiments the expression level of the expression product at a first time point is compared to the expression level of the same expression product at a second time point,

wherein an increase or decrease in expression at the second time point relative to the first time point is indicative of the progression of cancer.

[0054] The invention also provides methods for treating cancer in a patient comprising modulating a cancer-associated gene-activity. In some embodiments the cancer-associated gene-activity is cell proliferation, cell growth, cell motility, metastasis, cell migration, cell survival, or tumorigenicity. In some embodiments the methods comprise administering to the patient an antibody, a nucleic acid, or a polypeptide that inhibits the cancer-associated gene activity. In some embodiments the antibody is a neutralizing antibody. In some embodiments the antibody is a monoclonal antibody. In some embodiments the monoclonal antibody binds to a cancer-associated polypeptide with an affinity of at least $1 \times 10^8 K_a$. In some embodiments the monoclonal antibody inhibits one or more of cancer cell growth, tumor formation, cell survival and cancer cell proliferation. In some embodiments the antibody is a monoclonal antibody, a polyclonal antibody, a chimeric antibody, a human antibody, a humanized antibody, a single-chain antibody, a bi-specific antibody, a multi-specific antibody, or a Fab fragment.

[0055] The invention also provides methods of treating a cancer in a patient characterized by overexpression of a cancer-associated gene relative to a control. In some embodiments the methods comprise modulating a cancer-associated gene activity in the patient. In some embodiments the cancer-associated gene activity is selected from the group consisting of cell proliferation, cell growth, cell motility, metastasis, cell migration, cell survival, gene expression and tumorigenicity. In some embodiments the cancer is selected from the group consisting of cervical cancer, kidney cancer, ovarian cancer, pancreatic cancer and skin cancer. In some embodiments the methods comprise administering to the patient an antibody, a nucleic acid, or a polypeptide that inhibits the cancer-associated gene activity.

[0056] The present invention also provides methods for identifying a patient as susceptible to treatment with an antibody that binds to an expression product of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLA2, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFB3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA,

comprising measuring the level of the expression product of the gene in a biological sample from that patient.

[0057] The invention also provides compositions for treating, diagnosing or detecting cancer. In some embodiments the compositions comprise an antibody or oligonucleotide specific for an expression product of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA. In some embodiments the compositions further comprise a conventional cancer medicament. In some embodiments the compositions are pharmaceutical compositions. In some embodiments the compositions are sterile injectables.

[0058] The invention further provides assays for identifying a candidate agent that modulates the growth of a cancerous cell, comprising a) detecting the level of expression of one or more (i.e. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) expression product of a cancer-associated gene as listed in any of the above-described embodiments of the invention in the presence of the candidate agent; and b) comparing that level of expression with the level of expression in the absence of the candidate agent, wherein a difference in expression indicates that the candidate agent modulates the level of expression of the expression product of the cancer-associated gene.

[0059] The invention also provides methods for identifying an agent that modifies the expression level of a cancer-associated gene, comprising: a) contacting a cell expressing a cancer-associated gene as listed in any of the above-described embodiments of the invention with a candidate agent, and b) determining the effect of the candidate agent on the cell, wherein a change in expression level indicates that the candidate agent is able to modulate expression.

[0060] In some embodiments the candidate agent is a polynucleotide, a polypeptide, an antibody or a small organic molecule.

[0061] The invention also provides methods for detecting breast cancer in a biological sample comprising determining the sequence or expression level of one or more of PRLR, FGFR3, SELL, CD28, CDH23, FGFR1, PLAU, LTBR, ITGA9, TSPAN-2, CD2, IGSF8, ITGA4, PAG,

THY1, FBN2, LTB and COL19A1, which, among other genes of the invention, are correlated to breast cancer.

[0062] The invention also provides methods for detecting colon cancer in a biological sample comprising determining the sequence or expression level of one or more of NOTCH1, IL2RA, SELL, CD28, PDCD1LG2, LFNG, MET, PLAU, SLC4A8, LTBR, FGFR2, LNPEP, FGF8 and THY1, which, among other genes of the invention, are correlated to colon cancer.

[0063] The invention also provides methods for detecting kidney cancer in a biological sample comprising determining the sequence or expression level of one or more of IL2RA, FLT3, PDCD1LG2, LFNG, CDH23, FCMD, UTS2, XYLT1, SLC9A9 and PRG1, which, among other genes of the invention, are correlated to kidney cancer.

[0064] The invention also provides methods for detecting liver cancer in a biological sample comprising determining the sequence or expression level of one or more of WNT3A, FGFR1, COL15A1, IGSF8, TNFRSF14, THY1, SLC9A9, SLC12A7, FGF10 and TNFRSF8, which, among other genes of the invention, are correlated to liver cancer.

[0065] The invention also provides methods for detecting lung cancer in a biological sample comprising determining the sequence or expression level of one or more of PRLR, PTPRK, PLAU, ITGA9, COL15A1, TSPAN-2, LNPEP, THY1, IL7R, SLC1A2 and PDGFRA, which, among other genes of the invention, are correlated to lung cancer.

[0066] The invention also provides methods for detecting lymphoid cancer in a biological sample comprising determining the sequence or expression level of one or more of WNT3, IL2RA, TNFRSF6, TPCN1, IL4R, SLAMF6, CD74, CD2, LNPEP, PAG, FGF8, XYLT1, LTB, IL12A, TNFRSF8 and CD38, which, among other genes of the invention, are correlated to lymphoid cancer.

[0067] The invention also provides methods for detecting ovarian cancer in a biological sample comprising determining the sequence or expression level of one or more of WNT1, FGFR3, CDH23, IFNAR2, TPCN1, LTBR, SLAMF6, CD74, THSD2, IGSF8, UTS2, THY1, LTB, KCNH5, SLC12A7, FGF10, IL7R, SLC1A2, SLPI, IL12A, IL6R, CD38 and PDGFRA, which, among other genes of the invention, are correlated to ovary cancer.

[0068] The invention also provides methods for detecting pancreatic cancer in a biological sample comprising determining the sequence or expression level of one or more of PLAU, FCMD, IGSF8, THY1, FBN2, KCNH5, FGF10 and SLC1A2, which, among other genes of the invention, are correlated to pancreas cancer.

[0069] The invention also provides methods for detecting prostate cancer in a biological sample comprising determining the sequence or expression level of one or more of WNT3, WNT1, IL2RA, PRLR, TNFRSF6, SELL, CD28, PDCD1LG2, CDH23, FGFR1, IGSF8, TNFRSF14, TGFBR3, MATN2 and TNFRSF8, which, among other genes of the invention, are correlated to prostate cancer.

[0070] The invention also provides methods for detecting uterine cancer in a biological sample comprising determining the sequence or expression level of one or more of NOTCH1, IL2RA, TNFRSF6, CD28, FGFR1, TPCN1, ITGA9, COL15A1, CD74, TSPAN-2, PAG, TGFBR3, THY1, SLC12A7, MATN2, COL19A1 and SLC28A2, which, among other genes of the invention, are correlated to uterine cancer.

[0071] The invention also provides methods for detecting cervical cancer in a biological sample comprising determining the sequence or expression level of one or more of TNFRSF6, SELL, CD28, IFNAR2, PLAUI, FCMD, SLAMF6, CPM, SLC1A2, TNFRSF8 and CD38, which, among other genes of the invention, are correlated to cervical cancer.

[0072] The invention also provides methods for detecting stomach cancer in a biological sample comprising determining the sequence or expression level of one or more of LFNG, FGFR2, SLAMF6, PAG, KCNH5 and SLC1A2, which, among other genes of the invention, are correlated to stomach cancer.

[0073] The invention also provides methods for detecting skin cancer in a biological sample comprising determining the sequence or expression level of one or more of IL2RA, PRLR, CDH23, IFNAR2, IGSF8, TNFRSF14, KCNH5, SLC1A2 and TNFRSF8, which, among other genes of the invention, are correlated to skin cancer.

Definitions

[0074] The present invention identifies genes which are related to cancer (e.g. “cancer-associated genes”). Thus, polypeptides encoded by these genes are referred to as “cancer-associated polypeptides” or “cancer-associated proteins”. Nucleic acid sequences that encode these cancer-associated polypeptides are referred to as “cancer-associated polynucleotides”. Cells which encode and/or express a cancer-associated gene are referred to as “cancer-associated cells”. Cells which encode a cancer-associated gene are said to have a “cancer-associated genotype”. Cells which express a cancer-associated protein are said to have a “cancer-associated phenotype”. “Cancer-associated sequences” refers to both polypeptide and polynucleotide sequences derived from cancer-associated genes. “Cancer-associated nucleic acids” includes the

DNA comprising the cancer-associated gene, as well as mRNA and cDNA derived from that gene.

[0075] "Associated" in this context means that the nucleotide or protein sequences are either differentially expressed, activated, inactivated or altered in cancers as compared to normal tissue. As outlined below, cancer-associated sequences include those that are up-regulated (i.e. expressed at a higher level), as well as those that are down-regulated (i.e. expressed at a lower level), in cancers. Cancer-associated sequences also include sequences that have been altered (i.e., truncated sequences or sequences with substitutions, deletions or insertions, including point mutations) and show either the same expression profile or an altered profile. Generally, the cancer-associated sequences are from humans; however, as will be appreciated by those in the art, cancer-associated sequences from other organisms may be useful in animal models of disease and drug evaluation; thus, other cancer-associated sequences may be identified, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, and farm animals (including sheep, goats, pigs, cows, horses, etc). In some cases, prokaryotic cancer-associated sequences may be useful. Cancer-associated sequences from other organisms may be obtained using the techniques outlined below.

[0076] Cancer-associated sequences include recombinant nucleic acids. By the term "recombinant nucleic acid" herein is meant nucleic acid, originally formed in vitro, in general, by the manipulation of nucleic acid by polymerases and endonucleases, in a form not normally found in nature. Thus a recombinant nucleic acid is also an isolated nucleic acid, in a linear form, or cloned in a vector formed in vitro by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate using the in vivo cellular machinery of the host cell rather than in vitro manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated in vivo, are still considered recombinant or isolated for the purposes of the invention. As used herein a "polynucleotide" or "nucleic acid" is a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule. Thus, this term includes double- and single-stranded DNA and RNA. It also includes known types of modifications, for example, labels which are known in the art, methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, for

example proteins (including e.g., nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), those with intercalators (e.g., acridine, psoralen, etc.), those containing chelators (e.g., metals, radioactive metals, etc.), those containing alkylators, those with modified linkages (e.g., alpha anomeric nucleic acids, etc.), as well as unmodified forms of the polynucleotide.

[0077] As used herein, a polynucleotide "derived from" a designated sequence refers to a polynucleotide sequence which is comprised of a sequence of approximately at least about 6 nucleotides, at least about 8 nucleotides, at least about 10-12 nucleotides, and at least about 15-20 nucleotides corresponding to a region of the designated nucleotide sequence.

"Corresponding" means homologous to or complementary to the designated sequence. In some embodiments, the sequence of the region from which the polynucleotide is derived is homologous to or complementary to a sequence that is unique to a cancer-associated gene.

[0078] A "recombinant protein" is a protein made using recombinant techniques, i.e. through the expression of a recombinant nucleic acid as depicted above. A recombinant protein is distinguished from naturally occurring protein by at least one or more characteristics. For example, the protein may be isolated or purified away from some or all of the proteins and compounds with which it is normally associated in its wild type host, and thus may be substantially pure. For example, an isolated protein is unaccompanied by at least some of the material with which it is normally associated in its natural state, constituting at least about 0.5%, or at least about 5% by weight of the total protein in a given sample. A substantially pure protein comprises about 50-75%, at least about 80%, or at least about 90% by weight of the total protein. The definition includes the production of a cancer-associated protein from one organism in a different organism or host cell. Alternatively, the protein may be made at a significantly higher concentration than is normally seen, through the use of an inducible promoter or high expression promoter, such that the protein is made at increased concentration levels. Alternatively, the protein may be in a form not normally found in nature, as in the addition of an epitope tag or amino acid substitutions, insertions and deletions, as discussed below.

[0079] As used herein, the term "tag," "sequence tag" or "primer tag sequence" refers to an oligonucleotide with specific nucleic acid sequence that serves to identify a batch of polynucleotides bearing such tags therein. Polynucleotides from the same biological source are covalently tagged with a specific sequence tag so that in subsequent analysis the polynucleotide can be identified according to its source of origin. The sequence tags also serve as primers for nucleic acid amplification reactions.

[0080] A “microarray” is a linear or two-dimensional array of preferably discrete regions, each having a defined area, formed on the surface of a solid support. The density of the discrete regions on a microarray is determined by the total numbers of target polynucleotides to be detected on the surface of a single solid phase support, preferably at least about 50/cm², more preferably at least about 100/cm², even more preferably at least about 500/cm², and still more preferably at least about 1,000/cm². As used herein, a DNA microarray is an array of oligonucleotide primers placed on a chip or other surfaces used to amplify or clone target polynucleotides. Since the position of each particular group of primers in the array is known, the identities of the target polynucleotides can be determined based on their binding to a particular position in the microarray.

[0081] A “linker” is a synthetic oligodeoxyribonucleotide that contains a restriction site. A linker may be blunt end-ligated onto the ends of DNA fragments to create restriction sites that can be used in the subsequent cloning of the fragment into a vector molecule.

[0082] The term “label” refers to a composition capable of producing a detectable signal indicative of the presence of the target polynucleotide in an assay sample. Suitable labels include radioisotopes, nucleotide chromophores, enzymes, substrates, fluorescent molecules, chemiluminescent moieties, magnetic particles, bioluminescent moieties, and the like. As such, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical, chemical, or any other appropriate means. The term “label” is used to refer to any chemical group or moiety having a detectable physical property or any compound capable of causing a chemical group or moiety to exhibit a detectable physical property, such as an enzyme that catalyzes conversion of a substrate into a detectable product. The term “label” also encompasses compounds that inhibit the expression of a particular physical property. The label may also be a compound that is a member of a binding pair, the other member of which bears a detectable physical property.

[0083] The term “support” refers to conventional supports such as beads, particles, dipsticks, fibers, filters, membranes, and silane or silicate supports such as glass slides.

[0084] The term “amplify” is used in the broad sense to mean creating an amplification product which may include, for example, additional target molecules, or target-like molecules or molecules complementary to the target molecule, which molecules are created by virtue of the presence of the target molecule in the sample. In the situation where the target is a nucleic acid,

an amplification product can be made enzymatically with DNA or RNA polymerases or reverse transcriptases.

[0085] As used herein, a “biological sample” refers to a sample of tissue or fluid isolated from an individual, including but not limited to, for example, blood, plasma, serum, spinal fluid, lymph fluid, skin, respiratory, intestinal and genitourinary tracts, tears, saliva, milk, cells (including but not limited to blood cells), tumors, organs, and also samples of in vitro cell culture constituents..

[0086] The term “biological sources” as used herein refers to the sources from which the target polynucleotides are derived. The source can be of any form of “sample” as described above, including but not limited to, cell, tissue or fluid. “Different biological sources” can refer to different cells/tissues/organs of the same individual, or cells/tissues/organs from different individuals of the same species, or cells/tissues/organs from different species.

Cancer-associated genes

[0087] By “WNT3” we mean the gene “wingless-type MMTV integration site family, member 3” referred to by gene locus ID 7473 in the NCBI public database, having an mRNA referred to under accession number NM_030753 and encoding the polypeptide referred to under accession number NP_110380. WNT3 is a secreted protein. This gene underwent types I and II integration of the MMTV provirus in 72 cases. This gene was found to be overexpressed in 2 cancer tissue types tested, and to be overexpressed in 60% of lymphatic tumors sampled and in 32% of prostate tumors sampled. This means that this gene is correlated to lymphatic and prostate cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[0088] By “NOTCH1” we mean the gene “Notch homolog 1, translocation-associated” referred to by gene locus ID 4851 in the NCBI public database, having an mRNA referred to under accession number NM_017617 and encoding the polypeptide referred to under accession number NP_060087. NOTCH1 is a signalling receptor protein. This gene underwent types I and II integration of either the MLV or MMTV provirus in 132 cases. This gene was found to be overexpressed in 2 cancer tissue types tested, and to be overexpressed in 62% of colon tumors sampled and in 39% of uterine tumors sampled. This means that this gene is correlated to colon and uterine cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[0089] By “WNT1” we mean the gene “wingless-type MMTV integration site family, member 1” referred to by gene locus ID 7471 in the NCBI public database, having an mRNA referred to under accession number NM_005430 and encoding the polypeptide referred to under accession number NP_005421. WNT1 is a secreted protein. This gene underwent types I and II integration of the MLV and MMTV provirus in 346 cases. This gene was found to be overexpressed in 2 cancer tumors types tested, and to be overexpressed in 57% of ovarian tumors sampled. This gene was also found to be overexpressed in prostate tumors sampled (t-test). This means that this gene is correlated to ovarian cancer and prostate cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[0090] By “WNT3A” we mean the gene “wingless-type MMTV integration site family, member 3A” referred to by gene locus ID 89780 in the NCBI public database, having an mRNA referred to under accession number NM_033131 and encoding the polypeptide referred to under accession number NP_149122. WNT3A is a secreted protein. This gene underwent types I and II integration of the MLV or MMTV provirus in 106 cases. This gene was found to be overexpressed in 47% of liver tumors sampled. This means that this gene is correlated to liver cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[0091] By “IL2RA” we mean the gene “interleukin 2 receptor, alpha chain precursor” referred to by gene locus ID 3559 in the NCBI public database, having an mRNA referred to under accession number NM_000417 and encoding the polypeptide referred to under accession number NP_000408. IL2RA is a cytokine receptor. This gene underwent types I and II integration of either the MLV or MMTV provirus in 12 cases. This gene was found to be overexpressed in 6 of the cancer tissue types tested, and to be overexpressed in 32% of colon tumors sampled and in 28% of lymphatic tumors sampled. This gene was also found to be overexpressed in kidney, prostate, skin and uterine tumors sampled (t-test). This means that this gene is correlated to colon, lymphatic, prostate, skin and uterine cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[0092] By “FLT3” we mean the gene “fms-related tyrosine kinase 3” referred to by gene locus ID 2322 in the NCBI public database, having an mRNA referred to under accession number NM_004119 and encoding the polypeptide referred to under accession number NP_004110. FLT3 is a secreted protein. This gene underwent types I and II integration of the MLV provirus in 18 cases. This gene was found to be overexpressed in 36% of kidney tumors sampled. This

means that this gene is correlated to kidney cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[0093] By “PRLR” we mean the gene “prolactin receptor” referred to by gene locus ID 5618 in the NCBI public database, having an mRNA referred to under accession number NM_000949 and encoding the polypeptide referred to under accession number NP_000940. PRLR is a receptor protein. This gene underwent types I and II integration of either the MLV or MMTV provirus in 16 cases. This gene was found to be overexpressed in 4 of the cancer tissue types tested, and to be overexpressed in 45% of skin tumors sampled, 34% of lung tumors sampled and in 40% of breast tumors sampled. This gene was also found to be overexpressed in prostate tumors sampled (t-test). This means that this gene is correlated to skin, lung, breast and prostate cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[0094] By “TNFRSF6” we mean the gene “tumor necrosis factor receptor superfamily, member 6”, also known as “FAS” referred to by gene locus ID 355 in the NCBI public database, having an mRNA referred to under accession number NM_000043 and encoding the polypeptide referred to under accession number NP_000034. Transcriptional variants are known in the art and are referred to by mRNA accession nos. NM_152871, NM_152872, NM_152873, NM_152874, NM_152875, NM_152876, NM_152877 and protein accession nos. NP_690610, NP_690611, NP_690612, NP_690613, NP_690614, NP_690615, NP_690616 respectively. TNFRSF6 is a cytokine receptor. This gene underwent type II integration of the MLV provirus in 7 cases. This gene was found to be overexpressed in 4 of the cancer tissue types tested, and to be overexpressed in 50% of lymphatic tumors sampled, 32% of uterine tumors sampled, 28% of prostate tumors sampled and in 27% of cervical tumors sampled. This means that this gene is correlated to lymphatic, uterine, prostate and cervical cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[0095] By “PTPRK” we mean the gene “protein tyrosine phosphatase, receptor type, K” referred to by gene locus ID 5796 in the NCBI public database, having an mRNA referred to under accession number NM_002844 and encoding the polypeptide referred to under accession number NP_002835. PTPRK is a tyrosine phosphatase receptor. This gene underwent type II integration of the MLV provirus in 7 cases. This gene was found to be overexpressed in 33% of lung tumors sampled. This means that this gene is correlated to lung cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[0096] By “TNFSF11” we mean the gene “tumor necrosis factor (ligand) superfamily, member 11” referred to by gene locus ID 8600 in the NCBI public database, having an mRNA referred to under accession number NM_003701 and encoding the polypeptide referred to under accession number NP_003692. Transcriptional variants are known in the art and are referred to by mRNA accession number NM_033012 and protein accession number NP_143026 respectively. TNFSF11 is a cytokine. This gene underwent type I integration of the MLV provirus in 2 cases. This gene is therefore useful in the diagnosis and therapy of cancer.

[0097] By “FGFR3” we mean the gene “fibroblast growth factor receptor 3” referred to by gene locus ID 2261 in the NCBI public database, having an mRNA referred to under accession number NM_000142 and encoding the polypeptide referred to under accession number NP_000133. Transcriptional variants are known in the art and are referred to by mRNA accession number NM_022965 and protein accession number NP_075254 respectively. TNFSF11 is a cytokine. This gene is implicated in achondroplasia and thanatophoric dwarfism. FGFR3 is a tyrosine kinase receptor. This gene underwent types I and II integration of either the MLV or MMTV provirus in 9 cases. This gene was found to be overexpressed in 2 of the cancer tissue types tested, and to be overexpressed in 86% of ovarian tumors sampled and in 48% of breast tumors sampled. This means that this gene is correlated to ovarian and breast cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[0098] By “SELL” we mean the gene “selectin L (lymphocyte adhesion molecule 1)” referred to by gene locus ID 6402 in the NCBI public database, having an mRNA referred to under accession number NM_000655 and encoding the polypeptide referred to under accession number NP_000646. SELL is a secreted protein. This gene underwent type II integration of the MLV provirus in 7 cases. This gene was found to be overexpressed in 4 of the cancer tissue types tested, and to be overexpressed in 38% of breast tumors sampled and in 23% of cervical tumors sampled. This gene was also found to be overexpressed in colon and prostate tumors sampled (t-test). This means that this gene is correlated to breast, cervical, colon and prostate cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[0099] By “DPT” we mean the gene “dermatopontin” referred to by gene locus ID 1805 in the NCBI public database, having an mRNA referred to under accession number NM_001937 and encoding the polypeptide referred to under accession number NP_001928. DPT is an

extracellular matrix protein. This gene underwent type I integration of either the MLV or MMTV provirus in 3 cases. This gene is useful in the diagnosis and therapy of cancer.

[00100] By “CD28” we mean the gene “CD28 antigen (Tp44)” referred to by gene locus ID 940 in the NCBI public database, having an mRNA referred to under accession number NM_006139 and encoding the polypeptide referred to under accession number NP_006130. CD28 is a member of the Ig protein superfamily. This gene underwent types I and II integration of the MLV provirus in 4 cases. This gene was found to be overexpressed in 5 of the cancer tissue types tested, and to be overexpressed in 34% of uterine tumors sampled and 32% of colon tumors sampled. This gene was also found to be overexpressed in breast, cervical and prostate tumors sampled (t-test). This means that this gene is correlated to colon, uterine, breast, cervical and prostate cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00101] By “PDCD1LG2” we mean the gene “programmed cell death 1 ligand 2” referred to by gene locus ID 80380 in the NCBI public database, having an mRNA referred to under accession number NM_025239 and encoding the polypeptide referred to under accession number NP_079515. PDCD1LG2 is a secreted protein. This gene underwent types I and II integration of the MLV provirus in 7 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 36% of kidney tumors sampled. This gene was also found to be overexpressed in colon and prostate tumors sampled (t-test). This means that this gene is correlated to kidney, colon and prostate cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00102] By “LFNG” we mean the gene “lunatic fringe homolog (Drosophila)” referred to by gene locus ID 3955 in the NCBI public database, having an mRNA referred to under accession number NM_002304 and encoding the polypeptide referred to under accession number NP_002295. Transcriptional variants are known in the art and are referred to by mRNA accession nos. XM_496845, XM_499250 and protein accession nos. XP_496845, XP_499250 respectively. LFNG is a secreted protein. This gene underwent types I and II integration of the MLV provirus in 39 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 40% of colon tumors sampled, 32% of kidney tumors sampled and in 30% of stomach tumors sampled. This means that this gene is correlated to colon, kidney and stomach cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00103] By “ANXA2” we mean the gene “annexin A2” referred to by gene locus ID 302 in the NCBI public database, having an mRNA referred to under accession number NM_001002857 and encoding the polypeptide referred to under accession number NP_001002857.

Transcriptional variants are known in the art and are referred to by mRNA accession nos.

NM_001002858, NM_004039 and protein accession nos. NP_001002858, NP_004030 respectively. ANXA2 is a signalling protein. This gene underwent types I and II integration of either the MLV or MMTV provirus in 5 cases. This gene is useful in the diagnosis and therapy of cancer.

[00104] By “CDH23” we mean the gene “cadherin-like 23” referred to by gene locus ID 64072 in the NCBI public database, having an mRNA referred to under accession number NM_022124 and encoding the polypeptide referred to under accession number NP_071407. Transcriptional variants are known in the art and are referred to by mRNA accession number NM_052836 and protein accession number NP_443068 respectively. CDH23 is a member of the cadherin protein superfamily. This gene underwent type II integration of the MLV and MMTV provirus in 5 cases. This gene was found to be overexpressed in 5 of the cancer tissue types tested, and to be overexpressed in 70% of kidney tumors sampled, 38% of breast tumors sampled and in 36% of skin tumors sampled. This gene was also found to be overexpressed in ovarian and prostate tumors sampled (t-test). This means that this gene is correlated to kidney, breast, skin, ovarian and prostate cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00105] By “IFNAR2” we mean the gene “interferon (alpha, beta and omega) receptor 2” referred to by gene locus ID 3455 in the NCBI public database, having an mRNA referred to under accession number NM_000874 and encoding the polypeptide referred to under accession number NP_000865. Transcriptional variants are known in the art and are referred to by mRNA accession nos. NM_207584, NM_207585 and protein accession nos. NP_997467, NP_997468 respectively. IFNAR2 is a cytokine receptor. This gene underwent type II integration of either the MLV or MMTV provirus in 3 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 56% of skin tumors sampled and in 27% of cervical tumors sampled. This gene was also found to be overexpressed in ovarian tumors sampled (t-test). This means that this gene is correlated to skin, cervical and ovarian cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00106] By “RAMP1” we mean the gene “receptor (calcitonin) activity modifying protein 1” referred to by gene locus ID 10267 in the NCBI public database, having an mRNA referred to under accession number NM_005855 and encoding the polypeptide referred to under accession number NP_005846. RAMP1 is a receptor protein. This gene underwent type II integration of the MLV provirus in 11 cases. This gene is useful in the diagnosis and therapy of cancer.

[00107] By “MET”, as known as “MET tyrosine kinase receptor”, we mean the gene “met proto-oncogene (hepatocyte growth factor receptor)” referred to by gene locus ID 4233 in the NCBI public database, having an mRNA referred to under accession number NM_000245 and encoding the polypeptide referred to under accession number NP_000236.. This gene underwent type I integration of either the MLV or MMTV provirus in 7 cases. This gene was found to be overexpressed in 86% of colon tumors sampled. This means that this gene is correlated to colon cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[00108] By “FGFR1” we mean the gene “fibroblast growth factor receptor 1 (fms-related tyrosine kinase 2)” referred to by gene locus ID 2260 in the NCBI public database, having an mRNA referred to under accession number NM_000604 and encoding the polypeptide referred to under accession number NP_000595. Transcriptional variants are known in the art and are referred to by mRNA accession nos. NM_015850, NM_023105, NM_023106, NM_023107, NM_023108, NM_023109, NM_023110, NM_023111 and protein accession nos. NP_056934, NP_075593, NP_075594, NP_075595, NP_075596, NP_075597, NP_075598, NP_075599 respectively. This gene is implicated in Pfeiffer syndrome. FGFR1 is a tyrosine kinase receptor. This gene underwent types I and II integration of either the MLV or MMTV provirus in 9 cases. This gene was found to be overexpressed in 4 of the cancer tissue types tested, and to be overexpressed in 37% of uterine tumors sampled and in 30% of breast tumors sampled. This gene was also found to be overexpressed in liver prostate tumors sampled (t-test). This means that this gene is correlated to uterine, breast, liver and prostate cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers. -

[00109] By “PLAU” we mean the gene “plasminogen activator, urokinase” referred to by gene locus ID 5328 in the NCBI public database, having an mRNA referred to under accession number NM_002658 and encoding the polypeptide referred to under accession number NP_002649. PLAU is a receptor bound protease. This gene underwent type I integration of the MLV provirus 2 cases. This gene was found to be overexpressed in 95% of colon tumors sampled, 79% of lung tumors sampled, 65% of pancreatic tumors sampled and in 52% of breast

tumors sampled. This gene was also found to be overexpressed in cervical tumors sampled (t-test). This means that this gene is correlated to colon, lung, pancreatic, breast and cervical cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00110] By “SLC4A8” we mean the gene “solute carrier family 4, sodium bicarbonate cotransporter, member” referred to by gene locus ID 9498 in the NCBI public database, having an mRNA referred to under accession number NM_004858 and encoding the polypeptide referred to under accession number NP_004849. SLC4A8 is a carrier protein. This gene underwent types I and II integration of the MLV and MMTV provirus in 5 cases. This gene was found to be overexpressed in 50% of colon tumors sampled. This means that this gene is correlated to colon cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[00111] By “TPCN1” we mean the gene “two pore segment channel 1” referred to by gene locus ID 53373 in the NCBI public database, having an mRNA referred to under accession number NM_017901 and encoding the polypeptide referred to under accession number NP_060371. TPCN1 is an ion channel protein. This gene underwent type II integration of the MLV provirus in 3 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 29% of uterine tumors sampled and in 23% of lymphatic tumors sampled. This gene was also found to be overexpressed in ovarian tumors sampled (t-test). This means that this gene is correlated to uterine, lymphatic and ovarian cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00112] By “LTBR” we mean the gene “lymphotoxin beta receptor (TNFR superfamily, member 3)” referred to by gene locus ID 4055 in the NCBI public database, having an mRNA referred to under accession number NM_002342 and encoding the polypeptide referred to under accession number NP_002333. LTBR is a cytokine receptor. This gene underwent type I integration of the MLV and MMTV provirus in 5 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 40% of breast tumors sampled. This gene was also found to be overexpressed in colon and ovarian tumors sampled (t-test). This means that this gene is correlated to breast, colon and ovarian cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00113] By “IL4R” we mean the gene “interleukin 4 receptor” referred to by gene locus ID 3566 in the NCBI public database, having an mRNA referred to under accession number NM_000418 and encoding the polypeptide referred to under accession number NP_000409. Transcriptional variants are known in the art and are referred to by mRNA accession number NM_001008699 and protein accession number NP_001008699 respectively. IL4R is a cytokine receptor. This gene underwent type I integration of the MLV provirus in 4 cases. This gene was found to be overexpressed in 23% of lymphatic tumors sampled. This means that this gene is correlated to lymphatic cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[00114] By “ITGA9” we mean the gene “integrin, alpha 9” referred to by gene locus ID 3680 in the NCBI public database, having an mRNA referred to under accession number NM_002207 and encoding the polypeptide referred to under accession number NP_002198. ITGA9 is an integrin. This gene underwent type II integration of either the MLV or MMTV provirus in 3 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 38% of lung tumors sampled, 30% of breast tumors sampled and in 26% of uterine tumors sampled. This means that this gene is correlated to lung, breast and uterine cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00115] By “FGFR2” we mean the gene “fibroblast growth factor receptor 2 (bacteria-expressed kinase, keratinocyte growth factor receptor)” referred to by gene locus ID 2263 in the NCBI public database, having an mRNA referred to under accession number NM_000141 and encoding the polypeptide referred to under accession number NP_000132. Transcriptional variants are known in the art and are referred to by mRNA accession nos. NM_022969, NM_022970, NM_022971, NM_022972, NM_022973, NM_022974, NM_022975, NM_022976, NM_023028, NM_023029, NM_023030, NM_023031 and protein accession nos. NP_075258, NP_075259, NP_075260, NP_075261, NP_075262, NP_075263, NP_075264, NP_075265, NP_075417, NP_075418, NP_075419, NP_075420 respectively. This gene is implicated in craniofacial dysostosis 1, Crouzon syndrome, Pfeiffer syndrome and Jackson-Weiss syndrome. FGFR2 is a tyrosine kinase receptor. This gene underwent types I and II integration of the MMTV provirus in 94 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 40% of stomach tumors sampled and in 33% of colon tumors sampled. This means that this gene is correlated to colon and stomach

cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00116] By “FCMD” we mean the gene “Fukuyama type congenital muscular dystrophy (fukutin)” referred to by gene locus ID 2218 in the NCBI public database, having an mRNA referred to under accession number NM_006731 and encoding the polypeptide referred to under accession number NP_006722. FCMD is an extracellular matrix protein. This gene underwent type II integration of the MLV provirus in 2 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 33% of pancreatic tumors sampled and in 30% of cervical tumors sampled. This gene was also found to be overexpressed in kidney tumors sampled (t-test). This means that this gene is correlated to pancreatic, kidney and cervical cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00117] By “COL15A1” we mean the gene “collagen, type XV, alpha 1” referred to by gene locus ID 1306 in the NCBI public database, having an mRNA referred to under accession number NM_001855 and encoding the polypeptide referred to under accession number NP_001846. COL15A1 is a collagen type protein. This gene underwent types I and II integration of the MLV and MMTV provirus in 7 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 50% of liver tumors sampled, 38% of lung tumors sampled and in 29% of uterine tumors sampled. This means that this gene is correlated to liver, lung and uterine cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00118] By “SLAMF6” we mean the gene “SLAM family member 6” referred to by gene locus ID 114836 in the NCBI public database, having an mRNA referred to under accession number NM_052931 and encoding the polypeptide referred to under accession number NP_443163. SLAMF6 is a member of the Ig protein superfamily. This gene underwent type II integration of the MLV provirus in 4 cases. This gene was found to be overexpressed in 4 of the cancer tissue types tested, and to be overexpressed in 50% of ovarian tumors sampled, 43% lymphatic tumors sampled and in 30% of stomach tumors sampled. This gene was also found to be overexpressed in cervical tumors sampled (t-test). This means that this gene is correlated to ovarian, lymphatic, stomach and cervical cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00119] By “CD74” we mean the gene “CD74 antigen (invariant polypeptide of major histocompatibility complex, class II antigen-associated)” referred to by gene locus ID 972 in the NCBI public database, having an mRNA referred to under accession number NM_004355 and encoding the polypeptide referred to under accession number NP_004346. CD74 is a chaperone protein. This gene underwent types I and II integration of the MLV provirus in 36 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 26% of uterine tumors sampled and in 23% of lymphatic tumors sampled. This gene was also found to be overexpressed in ovarian tumors sampled (t-test). This means that this gene is correlated to uterine, lymphatic and ovarian cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00120] By “TSPAN-2” we mean the gene “tetraspan 2” referred to by gene locus ID 10100 in the NCBI public database, having an mRNA referred to under accession number NM_005725 and encoding the polypeptide referred to under accession number NP_005716. TSPAN-2 is a tetraspanin protein. This gene underwent type I integration of the MLV provirus in 2 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 45% of breast tumors sampled and in 26% of lung tumors sampled. This gene was also found to be overexpressed in uterine tumors sampled (t-test). This means that this gene is correlated to breast, lung and uterine cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00121] By “THSD2”, also known as “RSPO3” we mean the gene “thrombospondin, type I, domain containing 2” referred to by gene locus ID 84870 in the NCBI public database, having an mRNA referred to under accession number NM_032784 and encoding the polypeptide referred to under accession number NP_116173. THSD2 is an extracellular matrix protein. This gene underwent types I and II integration of the MMTV provirus in 20 cases. This gene was found to be overexpressed in 4 of the cancer tissue types tested, and to be overexpressed in 45% of breast tumors sampled, 43% of ovarian tumors sampled and 26% of uterine tumors sampled. This gene was also found to be overexpressed in pancreatic tumors sampled (t-test). This means that this gene is correlated to breast, ovarian, uterine and pancreatic cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00122] By “CD2” we mean the gene “CD2 antigen (p50), sheep red blood cell receptor” referred to by gene locus ID 914 in the NCBI public database, having an mRNA referred to under accession number NM_001767 and encoding the polypeptide referred to under accession number NP_001758. CD2 is a member of the Ig protein superfamily. This gene underwent type I

and II integration of the MLV and MMTV provirus in 5 cases. This gene was found to be overexpressed in 2 of the cancer tissue types tested, and to be overexpressed in 29% of breast tumors sampled and in 24% of lymphatic tumors sampled. This means that this gene is correlated to breast and lymphatic cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00123] By “LNPEP” we mean the gene “leucyl/cystinyl aminopeptidase” referred to by gene locus ID 4012 in the NCBI public database, having an mRNA referred to under accession number NM_005575 and encoding the polypeptide referred to under accession number NP_005566. LNPEP is an aminopeptidase. This gene underwent type I and II integration of the MLV provirus in 5 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 44% of colon tumors sampled, 42% of lung tumors sampled and in 40% of lymphatic tumors sampled. This means that this gene is correlated to colon, lung and lymphatic cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00124] By “IGSF8” we mean the gene “immunoglobulin superfamily, member 8” referred to by gene locus ID 93185 in the NCBI public database, having an mRNA referred to under accession number NM_052868 and encoding the polypeptide referred to under accession number NP_443100. IGSF8 is a member of the Ig protein superfamily. This gene underwent type II integration of the MLV provirus in 3 cases. This gene was found to be overexpressed in 6 of the cancer tissue types tested, and to be overexpressed in 65% of ovarian tumors sampled, 55% of liver tumors sampled, 52% of breast tumors sampled, 41% of prostate tumors sampled, 30% of pancreatic tumors sampled and in 30% of skin tumors sampled. This means that this gene is correlated to ovarian, liver, breast, prostate, pancreatic and skin cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00125] By “ATP1A2” we mean the gene “ATPase, Na⁺/K⁺ transporting, alpha 2 (+) polypeptide” referred to by gene locus ID 477 in the NCBI public database, having an mRNA referred to under accession number NM_000702 and encoding the polypeptide referred to under accession number NP_000693. ATP1A2 is a cation transporter ATPase. This gene underwent type II integration of the MLV provirus in 2 cases. This gene is therefore useful in the diagnosis and therapy of cancer.

[00126] By “ITGA4” we mean the gene “integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)” referred to by gene locus ID 3676 in the NCBI public database, having an

mRNA referred to under accession number NM_000885 and encoding the polypeptide referred to under accession number NP_000876. ITGA4 is an integrin. This gene underwent type I and II integration of the MLV and MMTV provirus in 6 cases. This gene was found to be overexpressed in 40% of breast tumors sampled. This means that this gene is correlated to breast cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[00127] By “B2M” we mean the gene “beta-2-microglobulin” referred to by gene locus ID 567 in the NCBI public database, having an mRNA referred to under accession number NM_004048 and encoding the polypeptide referred to under accession number NP_004039. B2M is a secreted protein. This gene underwent type I and II integration of the MLV provirus in 6 cases. This gene is useful in the diagnosis and therapy of cancer.

[00128] By “PAG”, also known as “PAG1”, we mean the gene “phosphoprotein associated with glycosphingolipid-enriched microdomains” referred to by gene locus ID 55824 in the NCBI public database, having an mRNA referred to under accession number NM_018440 and encoding the polypeptide referred to under accession number NP_060910. PAG is a receptor protein. This gene underwent type I and II integration of the MLV provirus in 6 cases. This gene was found to be overexpressed in 4 of the cancer tissue types tested, and to be overexpressed in 43% of stomach tumors sampled, 40% of uterine tumors sampled and in 33% of lymphatic tumors sampled. This gene was also found to be overexpressed in breast tumors sampled (t-test). This means that this gene is correlated to stomach, uterine, lymphatic and breast cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00129] By “NOTCH2” we mean the gene “Notch homolog 2” referred to by gene locus ID 4853 in the NCBI public database, having an mRNA referred to under accession number NM_024408 and encoding the polypeptide referred to under accession number NP_077719. NOTCH2 is a signalling receptor. This gene underwent type I and II integration of either the MLV or MMTV provirus in 40 cases. This gene is useful in the diagnosis and therapy of cancer.

[00130] By “UTS2” we mean the gene “urotensin 2” referred to by gene locus ID 10911 in the NCBI public database, having an mRNA referred to under accession number NM_006786 and encoding the polypeptide referred to under accession number NP_006777. Transcriptional variants are known in the art and are referred to by mRNA accession number NM_021995 and protein accession number NP_068835 respectively. This gene underwent type I integration of the MLV provirus in 2 cases. This gene was found to be overexpressed in 2 of the cancer tissue

types tested, and to be overexpressed in 37% of ovarian tumors sampled. This gene was also found to be overexpressed in kidney tumors sampled (t-test). This means that this gene is correlated to ovarian and kidney cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00131] By “TNFRSF14” we mean the gene “tumor necrosis factor receptor superfamily, member 14 (herpesvirus entry mediator)” referred to by gene locus ID 8764 in the NCBI public database, having an mRNA referred to under accession number NM_003820 and encoding the polypeptide referred to under accession number NP_003811. TNFRSF14 is a cytokine receptor. This gene underwent type I and II integration of the MLV provirus in 4 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 75% of skin tumors sampled, 48% of liver tumors sampled and in 36% of prostate tumors sampled. This means that this gene is correlated to skin, liver and prostate cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00132] By “TGFBR3” we mean the gene “transforming growth factor, beta receptor III (betaglycan, 300kDa)” referred to by gene locus ID 7049 in the NCBI public database, having an mRNA referred to under accession number NM_003243 and encoding the polypeptide referred to under accession number NP_003234. TGFBR3 is a receptor. This gene underwent type I and II integration of the MLV and MMTV provirus in 9 cases. This gene was found to be overexpressed in 2 of the cancer tissue types tested, and to be overexpressed in 67% of prostate tumors sampled and in 27% of uterine tumors sampled. This means that this gene is correlated to prostate and uterine cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00133] By “FGF8” we mean the gene “fibroblast growth factor 8 (androgen-induced)” referred to by gene locus ID 2253 in the NCBI public database, having an mRNA referred to under accession number NM_006119 and encoding the polypeptide referred to under accession number NP_006110. Transcriptional variants are known in the art and are referred to by mRNA accession nos. NM_033163, NM_033164, NM_033165 and protein accession nos. NP_149353, NP_149354, NP_149355 respectively. FGF8 is a secreted protein. This gene underwent type I and II integration of the MLV and MMTV provirus in 62 cases. This gene was found to be overexpressed in 2 of the cancer tissue types tested, and to be overexpressed in 33 % of lymphatic tumors sampled. This gene was also found to be overexpressed in colon tumors sampled (t-test). This means that this gene is correlated to lymphatic and colon cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00134] By “FGF3” we mean the gene “fibroblast growth factor 3 (murine mammary tumor virus integration site (v-int-2) oncogene homolog)” referred to by gene locus ID 2248 in the NCBI public database, having an mRNA referred to under accession number NM_005247 and encoding the polypeptide referred to under accession number NP_005238. FGF3 is a secreted protein. This gene underwent type I and II integration of the MLV and MMTV provirus in 824 cases. This gene is useful in the diagnosis and therapy of cancer.

[00135] By “XYLT1” we mean the gene “xylosyltransferase I” referred to by gene locus ID 64131 in the NCBI public database, having an mRNA referred to under accession number NM_022166 and encoding the polypeptide referred to under accession number NP_071449. XYLT1 is a secreted protein. This gene underwent type I and II integration of the MLV and MMTV provirus in 4 cases. This gene was found to be overexpressed in 2 of the cancer tissue types tested, and to be overexpressed in 40% of lymphatic tumors sampled and in 31% of kidney tumors sampled. This means that this gene is correlated to lymphatic and kidney cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00136] By “FGF4” we mean the gene “fibroblast growth factor 4 (heparin secretory transforming protein 1, Kaposi sarcoma oncogene)” referred to by gene locus ID 2249 in the NCBI public database, having an mRNA referred to under accession number NM_002007 and encoding the polypeptide referred to under accession number NP_001998. FGF4 is a secreted protein. This gene underwent type I and II integration of the MMTV provirus in 49 cases. This gene is useful in the diagnosis and therapy of cancer.

[00137] By “LOC390243” we mean the gene “LOC390243, similar to folate receptor 3” referred to by gene locus ID 390243 in the NCBI public database, having an mRNA referred to under accession number XM_372428 and encoding the polypeptide referred to under accession number XP_372428. LOC390243 is a receptor. This gene underwent type I and II integration of the MLV provirus in 2 cases. This gene is useful in the diagnosis and therapy of cancer.

[00138] By “THY1” we mean the gene “Thy-1 cell surface antigen” referred to by gene locus ID 7070 in the NCBI public database, having an mRNA referred to under accession number NM_006288 and encoding the polypeptide referred to under accession number NP_006279. THY1 is a member of the Ig protein superfamily. This gene underwent type I integration of the MLV provirus in 4 cases. This gene was found to be overexpressed in 7 of the cancer tissue types tested, and to be overexpressed in 59% of colon tumors sampled, 45% of liver tumors sampled, 50% of breast tumors sampled, 34% of lung tumors sampled, 26% of ovarian tumors

sampled and in 54% of uterine tumors sampled. This gene was also found to be overexpressed in pancreatic tumors sampled (t-test). This means that this gene is correlated to colon, liver, breast, lung, ovarian, uterine and pancreatic cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00139] By “SLC9A9” we mean the gene “solute carrier family 9 (sodium/hydrogen exchanger), isoform 9” referred to by gene locus ID 285195 in the NCBI public database, having an mRNA referred to under accession number NM_173653 and encoding the polypeptide referred to under accession number NP_775924. SLC9A9 is a carrier protein. This gene underwent type I and II integration of the MLV provirus in 4 cases. This gene was found to be overexpressed in 2 of the cancer tissue types tested, and to be overexpressed in 38% of kidney tumors sampled and in 33% of liver tumors sampled. This means that this gene is correlated to kidney and liver cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00140] By “PRG1” we mean the gene “proteoglycan 1, secretory granule” referred to by gene locus ID 5552 in the NCBI public database, having an mRNA referred to under accession number NM_002727 and encoding the polypeptide referred to under accession number NP_002718. PRG1 is a proteoglycan. This gene underwent type I and II integration of the MLV provirus in 2 cases. This gene was found to be overexpressed in lung tumors sampled (t-test). This means that this gene is correlated to lung cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[00141] By “CPM” we mean the gene “carboxypeptidase M” referred to by gene locus ID 1368 in the NCBI public database, having an mRNA referred to under accession number NM_001005502 and encoding the polypeptide referred to under accession number NP_001005502. Transcriptional variants are known in the art and are referred to by mRNA accession nos. NM_001874, NM_198320 and protein accession nos. NP_001865, NP_938079 respectively. CPM is a carboxypeptidase. This gene underwent type I and II integration of the MLV and MMTV provirus in 4 cases. This gene was found to be overexpressed in cervical tumors sampled (t-test). This means that this gene is correlated to cervical cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[00142] By “FBN2” we mean the gene “fibrillin 2” referred to by gene locus ID 2201 in the NCBI public database, having an mRNA referred to under accession number NM_001999 and encoding the polypeptide referred to under accession number NP_001990. This gene is

implicated in congenital contractural arachnodactyly. FBN2 is an extracellular matrix protein. This gene underwent type I and II integration of the MLV provirus in 2 cases. This gene was found to be overexpressed in 2 of the cancer tissue types tested, and to be overexpressed in 48% of breast tumors sampled and in 29% of pancreatic tumors sampled. This means that this gene is correlated to breast and pancreatic cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00143] By “LTB” we mean the gene “lymphotoxin beta (TNF superfamily, member 3)” referred to by gene locus ID 4050 in the NCBI public database, having an mRNA referred to under accession number NM_002341 and encoding the polypeptide referred to under accession number NP_002332. Transcriptional variants are known in the art and are referred to by mRNA accession number NM_009588 and protein accession number NP_033666 respectively. LTB is a cytokine. This gene underwent type II integration of the MLV provirus in 2 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 50% of ovarian tumors sampled, 29% of breast tumors sampled and in 27% of lymphatic tumors sampled. This means that this gene is correlated to ovarian, breast and lymphatic cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00144] By “KCNH5” we mean the gene “potassium voltage-gated channel, subfamily H (eag-related), member 5” referred to by gene locus ID 27133 in the NCBI public database, having an mRNA referred to under accession number NM_139318 and encoding the polypeptide referred to under accession number NP_647479. Transcriptional variants are known in the art and are referred to by mRNA accession nos. NM_172375, NM_172376 and protein accession nos. NP_758963, NP_758964 respectively. KCNH5 is potassium ion channel protein. This gene underwent type I and II integration of the MMTV provirus in 2 cases. This gene was found to be overexpressed in 4 of the cancer tissue types tested, and to be overexpressed in 60% of pancreatic tumors sampled, 46% of stomach tumors sampled, 44% of ovarian tumors sampled and in 36% of skin tumors sampled. This means that this gene is correlated to pancreatic, stomach, ovarian and skin cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00145] By “SLC12A7” we mean the gene “solute carrier family 12 (potassium/chloride transporters), member 7” referred to by gene locus ID 10723 in the NCBI public database, having an mRNA referred to under accession number NM_006598 and encoding the polypeptide referred to under accession number NP_006589. SLC12A7 is a carrier protein. This gene underwent type II integration of the MLV and MMYV provirus in 4 cases. This gene was found

to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 50% of uterine tumors sampled and in 48% of liver tumors sampled. This gene was also found to be overexpressed in ovarian tumors sampled (t-test). This means that this gene is correlated to uterine, liver and ovarian cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00146] By “FGF10” we mean the gene “fibroblast growth factor 10” referred to by gene locus ID 2255 in the NCBI public database, having an mRNA referred to under accession number NM_004465 and encoding the polypeptide referred to under accession number NP_004456. FGF10 is a secreted protein. This gene underwent type I and II integration of the MMTV provirus in 42 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 47% of liver tumors sampled. This gene was also found to be overexpressed in ovarian and pancreatic tumors sampled (t-test). This means that this gene is correlated to liver, ovarian and pancreatic cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00147] By “IL7R” we mean the gene “interleukin 7 receptor” referred to by gene locus ID 3575 in the NCBI public database, having an mRNA referred to under accession number NM_002185 and encoding the polypeptide referred to under accession number NP_002176. IL7R is a cytokine receptor. This gene underwent type I and II integration of the MLV provirus in 7 cases. This gene was found to be overexpressed in 2 of the cancer tissues tested, and to be overexpressed in lung tumors sampled and ovarian tumors sampled (t-test). This means that this gene is correlated to lung and ovarian cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00148] By “MATN2” we mean the gene “matrilin 2” referred to by gene locus ID 4147 in the NCBI public database, having an mRNA referred to under accession number NM_002380 and encoding the polypeptide referred to under accession number NP_002371. Transcriptional variants are known in the art and are referred to by mRNA accession number NM_030583 and protein accession number NP_085072 respectively. MATN2 is an extracellular matrix protein. This gene underwent type II integration of either the MLV or MMTV provirus in 2 cases. This gene was found to be overexpressed in 2 of the cancer tissue types tested, and to be overexpressed in 32% of prostate tumors sampled and in 24% of uterine tumors sampled. This means that this gene is correlated to prostate and uterine cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00149] By “COL19A1” we mean the gene “collagen, type XIX, alpha 1” referred to by gene locus ID 1310 in the NCBI public database, having an mRNA referred to under accession number NM_001858 and encoding the polypeptide referred to under accession number NP_001849. COL19A1 is an extracellular matrix protein. This gene underwent type II integration of either the MLV or MMTV provirus in 3 cases. This gene was found to be overexpressed in 2 of the cancer tissue types tested, and to be overexpressed in 27% of breast tumors sampled and in 24% of uterine tumors sampled. This means that this gene is correlated to breast and uterine cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00150] By “SLC1A2” we mean the gene “solute carrier family 1 (glial high affinity glutamate transporter), member 2” referred to by gene locus ID 6506 in the NCBI public database, having an mRNA referred to under accession number NM_004171 and encoding the polypeptide referred to under accession number NP_004162. SLC1A2 is a calcium channel transporter protein. This gene underwent type I and II integration of the MLV provirus in 2 cases. This gene was found to be overexpressed in 9 of the cancer tissue types tested, and to be overexpressed in 61% of ovarian tumors sampled, 60% of breast tumors sampled, 53% of lung tumors sampled, 40% of skin tumors sampled, 40% of lymphatic tumors sampled, 39% of prostate tumors sampled, 33% of liver tumors sampled and in 33% of uterine tumors sampled. This gene was also found to be overexpressed in pancreatic tumors sampled (t-test). This means that this gene is correlated to ovarian, breast, lung, skin, lymphatic, prostate, liver, uterine and pancreatic cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00151] By “SLC28A2” we mean the gene “solute carrier family 28 (sodium-coupled nucleoside transporter), member 2” referred to by gene locus ID 9153 in the NCBI public database, having an mRNA referred to under accession number NM_004212 and encoding the polypeptide referred to under accession number NP_004203. SLC28A2 is a carrier protein. This gene underwent type I integration of the MLV provirus in 2 cases. This gene was found to be overexpressed in 45% of uterine tumors sampled. This means that this gene is correlated to uterine cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[00152] By “SLPI” we mean the gene “secretory leukocyte protease inhibitor (antileukoproteinase)” referred to by gene locus ID 6590 in the NCBI public database, having an mRNA referred to under accession number NM_003064 and encoding the polypeptide referred

to under accession number NP_003055. SLPI is a secreted protein. This gene underwent type I and II integration of the MLV provirus in 5 cases. This gene was found to be overexpressed in 100% of ovarian tumors sampled. This means that this gene is correlated to ovarian cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[00153] By “IL12A” we mean the gene “interleukin 12A (natural killer cell stimulatory factor 1, cytotoxic lymphocyte maturation factor 1, p35)” referred to by gene locus ID 3592 in the NCBI public database, having an mRNA referred to under accession number NM_000882 and encoding the polypeptide referred to under accession number NP_000873. IL12A is a secreted protein. This gene underwent type I and II integration of the MLV provirus in 5 cases. This gene was found to be overexpressed in 2 of the cancer tissue types tested, and to be overexpressed in 47% of lymphatic tumors sampled. This means that this gene is correlated to lymphatic cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[00154] By “IL6R” we mean the gene “interleukin 6 receptor” referred to by gene locus ID 3570 in the NCBI public database, having an mRNA referred to under accession number NM_000565 and encoding the polypeptide referred to under accession number NP_000556. Transcriptional variants are known in the art and are referred to by mRNA accession number NM_181359 and protein accession number NP_852004 respectively. IL6R is a cytokine receptor. This gene underwent type I and II integration of the MLV provirus in 4 cases. This gene was found to be overexpressed in ovarian tumors sampled (t-test). This means that this gene is correlated to ovarian cancer, among other cancers, and is therefore a target for the diagnosis and therapy of this and other cancers.

[00155] By “TNFSF8” we mean the gene “tumor necrosis factor (ligand) superfamily, member 8” referred to by gene locus ID 944 in the NCBI public database, having an mRNA referred to under accession number NM_001244 and encoding the polypeptide referred to under accession number NP_001235. TNFSF8 is a cytokine. This gene underwent type I integration of the MLV provirus in 2 cases. This gene is useful in the diagnosis and therapy of cancer.

[00156] By “TNFRSF8” we mean the gene “tumor necrosis factor receptor superfamily, member 8” referred to by gene locus ID 943 in the NCBI public database, having an mRNA referred to under accession number NM_001243 and encoding the polypeptide referred to under accession number NP_001234. Transcriptional variants are known in the art and are referred to

by mRNA accession number NM_152942 and protein accession number NP_694421 respectively. TNFRSF8 is a cytokine receptor. This gene underwent type I and II integration of the MLV provirus in 5 cases. This gene was found to be overexpressed in 5 of the cancer tissue types tested, and to be overexpressed in 56% of liver tumors sampled, 45% of cervical tumors sampled, 44% of skin tumors sampled, 39% of prostate tumors sampled and in 33% of lymphatic tumors sampled. This means that this gene is correlated to liver, cervical, skin, prostate and lymphatic cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00157] By “CD38” we mean the gene “CD38 antigen (p45)” referred to by gene locus ID 952 in the NCBI public database, having an mRNA referred to under accession number NM_001775 and encoding the polypeptide referred to under accession number NP_001766. CD38 is a receptor. This gene underwent type II integration of either the MLV or MMTV provirus in 3 cases. This gene was found to be overexpressed in 3 of the cancer tissue types tested, and to be overexpressed in 35 % of ovarian tumors sampled and in 30% of lymphatic tumors sampled. This gene was also found to be overexpressed in cervical tumors sampled (t-test). This means that this gene is correlated to ovarian, lymphatic and cervical cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00158] By “PDGFRA” we mean the gene “platelet-derived growth factor receptor, alpha polypeptide” referred to by gene locus ID 5156 in the NCBI public database, having an mRNA referred to under accession number NM_006206 and encoding the polypeptide referred to under accession number NP_006197. PDGFRA is a tyrosine kinase receptor. This gene underwent type I and II integration of the MMTV provirus in 11 cases. This gene was found to be overexpressed in 2 of the cancer tissue types tested, and to be overexpressed in 27 % of lung tumors sampled. This gene was also found to be overexpressed in ovarian tumors sampled (t-test). This means that this gene is correlated to lung and ovarian cancer, among other cancers, and is therefore a target for the diagnosis and therapy of these and other cancers.

[00159] Therefore the presence or absence of expression of one of these genes alone may be sufficient to cause cancer. Alternatively an increase or decrease in expression of one of these genes (whether the expression of the gene is increased or decreased in cancer is dependent, in some embodiments on the gene and the cancer type) may be sufficient to cause cancer. In a further alternative, cancer may be induced when the expression of one or more (i.e. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) of these genes reaches or exceeds a threshold level. The threshold level

may be represented as a percentage increase or decrease in expression of the gene when compared with that in a "normal" control level of expression.

[00160] In some embodiments, differential expression or amplification of the genes of the invention may be evaluated using an in vivo diagnostic assay, e.g. by administering a molecule (such as an antibody) which binds the molecule to be detected and is tagged with a detectable label (e.g. a radioactive isotope) and externally scanning the patient for localization of the label.

[00161] In some embodiments, genes or gene expression products are selected for targeting by comparison of the expression level of the gene or gene expression product in comparison with neighboring healthy tissue or with pooled normal tissue. In some embodiments there is at least a 3 fold (300%) increased expression relative to normal tissue and/or control. In some embodiments a 3 fold increase is seen in comparison with a majority of pooled, commercially available normal tissue samples. Screening can also be carried out using laser capture microscopy to dissect cancerous tissues from normal adjacent ones, followed by expressional microarray analysis utilizing standard commercially available chips such as the standard Affimetrix chip U133 (cat# 900370) (see for example, Yang et al, (2005) *Oncogene*, 10-31). In some embodiments, custom chips containing nucleic acid samples derived from pools of patient tissue samples grouped by cancer type can be made and probed to analyze expression profiles (see for example Makino et al, *Dis Esophagus*. 2005;18(1):37-40.).

[00162] The invention also allows the use of homologs, fragments, and functional equivalents of the above-referenced cancer-associated genes. Homology can be based on the full gene sequence referenced above and is generally determined as outlined below, using homology programs or hybridization conditions. A homolog of a cancer-associated gene has preferably greater than about 75% (i.e. at least 80, at least 85, at least 90, at least 92, at least 94, at least 95, at least 96, at least 97, at least 98, at least 99% or more) homology with the cancer-associated gene. Such homologs may include splice variants, deletion, addition and/or substitution mutants and generally have functional similarity.

[00163] Homology in this context means sequence similarity or identity. One comparison for homology purposes is to compare the sequence containing sequencing errors to the correct sequence. This homology will be determined using standard techniques known in the art, including, but not limited to, the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *PNAS USA*

85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, WI), the Best Fit sequence program described by Devereux et al., Nucl. Acid Res. 12:387-395 (1984), in some embodiments using the default settings, or by inspection.

[00164] One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments. It can also plot a tree showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, J. Mol. Evol. 35:351-360 (1987); the method is similar to that described by Higgins & Sharp CABIOS 5:151-153 (1989). Useful PILEUP parameters include a default gap weight of 3.00, a default gap length weight of 0.10, and weighted end gaps.

[00165] Another example of a useful algorithm is the BLAST (Basic Local Alignment Search Tool) algorithm, described in Altschul et al., J. Mol. Biol. 215, 403-410, (1990) and Karlin et al., PNAS USA 90:5873-5787 (1993). A particularly useful BLAST program is the WU-BLAST-2 program which was obtained from Altschul et al., Methods in Enzymology, 266: 460-480 (1996); <http://blast.wustl.edu/>. WU-BLAST-2 uses several search parameters, most of which are set to the default values. The adjustable parameters are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11. The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched; however, the values may be adjusted to increase sensitivity. A percent amino acid sequence identity value is determined by the number of matching identical residues divided by the total number of residues of the "longer" sequence in the aligned region. The "longer" sequence is the one having the most actual residues in the aligned region (gaps introduced by WU-Blast-2 to maximize the alignment score are ignored).

[00166] The alignment may include the introduction of gaps in the sequences to be aligned. In addition, for sequences which contain either more or fewer nucleotides than those of the cancer-associated genes, it is understood that the percentage of homology will be determined based on the number of homologous nucleosides in relation to the total number of nucleosides. Thus homology of sequences shorter than those of the sequences identified herein will be determined using the number of nucleosides in the shorter sequence.

[00167] In some embodiments of the invention, polynucleotide compositions are provided that are capable of hybridizing under moderate to high stringency conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof.

Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5x SSC ("saline sodium citrate"; 9 mM NaCl, 0.9 mM sodium citrate), 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50-60° C, 5x SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2x, 0.5x and 0.2x SSC containing 0.1% SDS. One skilled in the art will understand that the stringency of hybridization can be readily manipulated, such as by altering the salt content of the hybridization solution and/or the temperature at which the hybridization is performed. For example, in some embodiments, suitable highly stringent hybridization conditions include those described above, with the exception that the temperature of hybridization is increased, e.g., to 60-65°C, or 65-70°C. Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide.

[00168] Thus nucleic acids that hybridize under high stringency to the nucleic acids identified throughout the present application and sequence listing, or their complements, are considered cancer-associated sequences. High stringency conditions are known in the art; see for example Maniatis et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edition, 1989, and *Short Protocols in Molecular Biology*, ed. Ausubel, et al., both of which are hereby incorporated by reference. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g. 10 to 50 nucleotides) and at least about 60°C for longer probes (e.g. greater than 50

nucleotides). In another embodiment, less stringent hybridization conditions are used; for example, moderate or low stringency conditions may be used, as are known in the art; see Maniatis and Ausubel, *supra*, and Tijssen, *supra*.

Detection of cancer-associated gene expression

[00169] The cancer-associated gene may be cloned and, if necessary, its constituent parts recombined to form the entire cancer-associated nucleic acid. Once isolated from its natural source, e.g., contained within a plasmid or other vector or excised therefrom as a linear nucleic acid segment, the recombinant cancer-associated nucleic acid can be further used as a probe to identify and isolate other cancer-associated nucleic acids, for example additional coding regions. It can also be used as a "precursor" nucleic acid to make modified or variant cancer-associated nucleic acids and proteins. The nucleotide sequence of the cancer-associated gene can also be used to design probes specific for the cancer-associated gene.

[00170] The cancer-associated nucleic acids may be used in several ways. Nucleic acid probes hybridizable to cancer-associated nucleic acids can be made and attached to biochips to be used in screening and diagnostic methods, or for gene therapy and/or antisense applications. Alternatively, the cancer-associated nucleic acids that include coding regions of cancer-associated proteins can be put into expression vectors for the expression of cancer-associated proteins, again either for screening purposes or for administration to a patient.

[00171] One such system for quantifying gene expression is kinetic polymerase chain reaction (PCR). Kinetic PCR allows for the simultaneous amplification and quantification of specific nucleic acid sequences. The specificity is derived from synthetic oligonucleotide primers designed to preferentially adhere to single-stranded nucleic acid sequences bracketing the target site. This pair of oligonucleotide primers forms specific, non-covalently bound complexes on each strand of the target sequence. These complexes facilitate in vitro transcription of double-stranded DNA in opposite orientations. Temperature cycling of the reaction mixture creates a continuous cycle of primer binding, transcription, and re-melting of the nucleic acid to individual strands. The result is an exponential increase of the target dsDNA product. This product can be quantified in real time either through the use of an intercalating dye or a sequence specific probe. SYBR® Greene I, is an example of an intercalating dye, that preferentially binds to dsDNA resulting in a concomitant increase in the fluorescent signal. Sequence specific probes, such as used with TaqMan® technology, consist of a fluorochrome and a quenching molecule covalently bound to opposite ends of an oligonucleotide. The probe is designed to selectively bind the target

DNA sequence between the two primers. When the DNA strands are synthesized during the PCR reaction, the fluorochrome is cleaved from the probe by the exonuclease activity of the polymerase resulting in signal dequenching. The probe signaling method can be more specific than the intercalating dye method, but in each case, signal strength is proportional to the dsDNA product produced. Each type of quantification method can be used in multi-well liquid phase arrays with each well representing primers and/or probes specific to nucleic acid sequences of interest. When used with messenger RNA preparations of tissues or cell lines, an array of probe/primer reactions can simultaneously quantify the expression of multiple gene products of interest. See Germer, S., et al., *Genome Res.* 10:258-266 (2000); Heid, C. A., et al., *Genome Res.* 6, 986-994 (1996).

[00172] Recent developments in DNA microarray technology make it possible to conduct a large scale assay of a plurality of target cancer-associated nucleic acid molecules on a single solid phase support. U.S. Pat. No. 5,837,832 (Chee et al.) and related patent applications describe immobilizing an array of oligonucleotide probes for hybridization and detection of specific nucleic acid sequences in a sample. Target polynucleotides of interest isolated from a tissue of interest are hybridized to the DNA chip and the specific sequences detected based on the target polynucleotides' preference and degree of hybridization at discrete probe locations. One important use of arrays is in the analysis of differential gene expression, where the profile of expression of genes in different cells, often a cell of interest and a control cell, is compared and any differences in gene expression among the respective cells are identified. Such information is useful for the identification of the types of genes expressed in a particular cell or tissue type and diagnosis of cancer conditions based on the expression profile.

[00173] Typically, RNA from the sample of interest is subjected to reverse transcription to obtain labeled cDNA. See U.S. Pat. No. 6,410,229 (Lockhart et al.) The cDNA is then hybridized to oligonucleotides or cDNAs of known sequence arrayed on a chip or other surface in a known order. The location of the oligonucleotide to which the labeled cDNA hybridizes provides sequence information on the cDNA, while the amount of labeled hybridized RNA or cDNA provides an estimate of the relative representation of the RNA or cDNA of interest. See Schena, et al. *Science* 270:467-470 (1995). For example, use of a cDNA microarray to analyze gene expression patterns in human cancer is described by DeRisi, et al. (*Nature Genetics* 14:457-460 (1996)).

[00174] Nucleic acid probes corresponding to cancer-associated nucleic acids may be made. Typically, these probes are synthesized based on the disclosed cancer-associated genes. The

nucleic acid probes attached to the biochip are designed to be substantially complementary to the cancer-associated nucleic acids, i.e. the target sequence (either the target sequence of the sample or to other probe sequences, for example in sandwich assays), such that specific hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect, in that there may be any number of base pair mismatches that will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. It is expected that the overall homology of the genes at the nucleotide level will be about 40% or greater, about 60% or greater, or about 80% or greater; and in addition that there will be corresponding contiguous sequences of about 8-12 nucleotides or longer. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by "substantially complementary" herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under normal reaction conditions, particularly high stringency conditions, as outlined herein. Whether or not a sequence is unique to a cancer-associated gene according to this invention can be determined by techniques known to those of skill in the art. For example, the sequence can be compared to sequences in databanks, e.g., GeneBank, to determine whether it is present in the uninfected host or other organisms. The sequence can also be compared to the known sequences of other viral agents, including those that are known to induce cancer.

[00175] A nucleic acid probe is generally single stranded but can be partly single and partly double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. In general, the oligonucleotide probes range from about 6, 8, 10, 12, 15, 20, 30 to about 100 bases long, from about 10 to about 80 bases, or from about 30 to about 50 bases. In some embodiments entire genes are used as probes. In some embodiments, much longer nucleic acids can be used, up to hundreds of bases. The probes are sufficiently specific to hybridize to complementary template sequence under conditions known by those of skill in the art. The number of mismatches between the probes sequences and their complementary template (target) sequences to which they hybridize during hybridization generally do not exceed 15%, 10% or 5%, as determined by FASTA (default settings).

[00176] Oligonucleotide probes can include the naturally-occurring heterocyclic bases normally found in nucleic acids (uracil, cytosine, thymine, adenine and guanine), as well as modified bases and base analogues. Any modified base or base analogue compatible with hybridization of the probe to a target sequence is useful in the practice of the invention. The sugar or glycoside

portion of the probe can comprise deoxyribose, ribose, and/or modified forms of these sugars, such as, for example, 2'-O-alkyl ribose. In some embodiments, the sugar moiety is 2'-deoxyribose; however, any sugar moiety that is compatible with the ability of the probe to hybridize to a target sequence can be used.

[00177] The nucleoside units of the probe may be linked by a phosphodiester backbone, as is well known in the art. In some embodiments, internucleotide linkages can include any linkage known to one of skill in the art that is compatible with specific hybridization of the probe including, but not limited to phosphorothioate, methylphosphonate, sulfamate (e.g., U.S. Patent No. 5,470,967) and polyamide (i.e., peptide nucleic acids). Peptide nucleic acids are described in Nielsen et al. (1991) *Science* 254: 1497-1500, U.S. Patent No. 5,714,331, and Nielsen (1999) *Curr. Opin. Biotechnol.* 10:71-75.

[00178] The probe can be a chimeric molecule; i.e., can comprise more than one type of base or sugar subunit, and/or the linkages can be of more than one type within the same primer. The probe can comprise a moiety to facilitate hybridization to its target sequence, as are known in the art, for example, intercalators and/or minor groove binders. Variations of the bases, sugars, and internucleoside backbone, as well as the presence of any pendant group on the probe, will be compatible with the ability of the probe to bind, in a sequence-specific fashion, with its target sequence. A large number of structural modifications, both known and to be developed, are possible within these bounds. Advantageously, the probes according to the present invention may have structural characteristics such that they allow the signal amplification, such structural characteristics being, for example, branched DNA probes as those described by Urdea et al. (*Nucleic Acids Symp. Ser.*, 24:197-200 (1991)) or in the European Patent No. EP-0225,807. Moreover, synthetic methods for preparing the various heterocyclic bases, sugars, nucleosides and nucleotides that form the probe, and preparation of oligonucleotides of specific predetermined sequence, are well-developed and known in the art. A method for oligonucleotide synthesis incorporates the teaching of U.S. Patent No. 5,419,966.

[00179] Multiple probes may be designed for a particular target nucleic acid to account for polymorphism and/or secondary structure in the target nucleic acid, redundancy of data and the like. In some embodiments, where more than one probe per sequence is used, either overlapping probes or probes to different sections of a single target cancer-associated gene are used. That is, two, three, four or more probes, with three being preferred, are used to build in a redundancy for a particular target. The probes can be overlapping (i.e. have some sequence in common), or specific for distinct sequences of a cancer-associated gene. When multiple target polynucleotides

are to be detected according to the present invention, each probe or probe group corresponding to a particular target polynucleotide is situated in a discrete area of the microarray.

[00180] Probes may be in solution, such as in wells or on the surface of a micro-array, or attached to a solid support. Examples of solid support materials that can be used include a plastic, a ceramic, a metal, a resin, a gel and a membrane. Useful types of solid supports include plates, beads, magnetic material, microbeads, hybridization chips, membranes, crystals, ceramics and self-assembling monolayers. Some embodiments comprise a two-dimensional or three-dimensional matrix, such as a gel or hybridization chip with multiple probe binding sites (Pevzner et al., *J. Biomol. Struct. & Dyn.* 9:399-410, 1991; Maskos and Southern, *Nuc. Acids Res.* 20:1679-84, 1992). Hybridization chips can be used to construct very large probe arrays that are subsequently hybridized with a target nucleic acid. Analysis of the hybridization pattern of the chip can assist in the identification of the target nucleotide sequence. Patterns can be manually or computer analyzed, but it is clear that positional sequencing by hybridization lends itself to computer analysis and automation. Algorithms and software, which have been developed for sequence reconstruction, are applicable to the methods described herein (R. Drmanac et al., *J. Biomol. Struct. & Dyn.* 5:1085-1102, 1991; P. A. Pevzner, *J. Biomol. Struct. & Dyn.* 7:63-73, 1989).

[00181] As will be appreciated by those in the art, nucleic acids can be attached or immobilized to a solid support in a wide variety of ways. By "immobilized" herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can be covalent or non-covalent. By "non-covalent binding" and grammatical equivalents herein is meant one or more of either electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, such as streptavidin, to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By "covalent binding" and grammatical equivalents herein is meant that the two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions.

[00182] Nucleic acid probes may be attached to the solid support by covalent binding such as by conjugation with a coupling agent or by, covalent or non-covalent binding such as

electrostatic interactions, hydrogen bonds or antibody-antigen coupling, or by combinations thereof. Typical coupling agents include biotin/avidin, biotin/streptavidin, *Staphylococcus aureus* protein A/IgG antibody Fc fragment, and streptavidin/protein A chimeras (T. Sano and C. R. Cantor, *Bio/Technology* 9:1378-81 (1991)), or derivatives or combinations of these agents. Nucleic acids may be attached to the solid support by a photocleavable bond, an electrostatic bond, a disulfide bond, a peptide bond, a diester bond or a combination of these sorts of bonds. The array may also be attached to the solid support by a selectively releasable bond such as 4,4'-dimethoxytrityl or its derivative. Derivatives which have been found to be useful include 3 or 4 [bis-(4-methoxyphenyl)]-methyl-benzoic acid, N-succinimidyl-3 or 4 [bis-(4-methoxyphenyl)]-methyl-benzoic acid, N-succinimidyl-3 or 4 [bis-(4-methoxyphenyl)]-hydroxymethyl-benzoic acid, N-succinimidyl-3 or 4 [bis-(4-methoxyphenyl)]-chloromethyl-benzoic acid, and salts of these acids.

[00183] Probes may be attached to biochips in a wide variety of ways, as will be appreciated by those in the art. As described herein, the nucleic acids can either be synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

[00184] Biochips comprise a suitable solid substrate. By "substrate" or "solid support" or other grammatical equivalents herein is meant any material that can be modified to contain discrete individual sites appropriate for the attachment or association of the nucleic acid probes and is amenable to at least one detection method. The solid phase support of the present invention can be of any solid materials and structures suitable for supporting nucleotide hybridization and synthesis. Preferably, the solid phase support comprises at least one substantially rigid surface on which the primers can be immobilized and the reverse transcriptase reaction performed. The substrates with which the polynucleotide microarray elements are stably associated may be fabricated from a variety of materials, including plastics, ceramics, metals, acrylamide, cellulose, nitrocellulose, glass, polystyrene, polyethylene vinyl acetate, polypropylene, polymethacrylate, polyethylene, polyethylene oxide, polysilicates, polycarbonates, Teflon®, fluorocarbons, nylon, silicon rubber, polyanhydrides, polyglycolic acid, polylactic acid, polyorthoesters, polypropylfumerate, collagen, glycosaminoglycans, and polyamino acids. Substrates may be two-dimensional or three-dimensional in form, such as gels, membranes, thin films, glasses, plates, cylinders, beads, magnetic beads, optical fibers, woven fibers, etc. One form of array is a three-dimensional array. One type of three-dimensional array is a collection of tagged beads. Each tagged bead has different primers attached to it. Tags are detectable by signaling means such as color (Luminex, Illumina) and electromagnetic field (Pharmaseq) and signals on tagged

beads can even be remotely detected (e.g., using optical fibers). The size of the solid support can be any of the standard microarray sizes, useful for DNA microarray technology, and the size may be tailored to fit the particular machine being used to conduct a reaction of the invention. In general, the substrates allow optical detection and do not appreciably fluoresce.

[00185] The surface of the biochip and the probe may be derivatized with chemical functional groups for subsequent attachment of the two. Thus, for example, the biochip is derivatized with a chemical functional group including, but not limited to, amino groups, carboxy groups, oxo groups and thiol groups, with amino groups being particularly preferred. Using these functional groups, the probes can be attached using functional groups on the probes. For example, nucleic acids containing amino groups can be attached to surfaces comprising amino groups, for example using linkers as are known in the art; for example, homo- or hetero-bifunctional linkers as are well known (see 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200, incorporated herein by reference). In addition, in some cases, additional linkers, such as alkyl groups (including substituted and heteroalkyl groups) may be used.

[00186] The oligonucleotides may be synthesized as is known in the art, and then attached to the surface of the solid support. As will be appreciated by those skilled in the art, either the 5' or 3' terminus may be attached to the solid support, or attachment may be via an internal nucleoside. In an additional embodiment, the immobilization to the solid support may be very strong, yet non-covalent. For example, biotinylated oligonucleotides can be made, which bind to surfaces covalently coated with streptavidin, resulting in attachment.

[00187] Arrays may be produced according to any convenient methodology, such as preforming the polynucleotide microarray elements and then stably associating them with the surface. Alternatively, the oligonucleotides may be synthesized on the surface, as is known in the art. A number of different array configurations and methods for their production are known to those of skill in the art and disclosed in WO 95/25116 and WO 95/35505 (photolithographic techniques), U.S. Patent Number 5,445,934 (in situ synthesis by photolithography), U.S. Patent Number 5,384,261 (in situ synthesis by mechanically directed flow paths); and U.S. Patent Number 5,700,637 (synthesis by spotting, printing or coupling); the disclosure of which are herein incorporated in their entirety by reference. Another method for coupling DNA to beads uses specific ligands attached to the end of the DNA to link to ligand-binding molecules attached to a bead. Possible ligand-binding partner pairs include biotin-avidin/streptavidin, or various antibody/antigen pairs such as digoxigenin-antidigoxigenin antibody (Smith *et al.*, "Direct

Mechanical Measurements of the Elasticity of Single DNA Molecules by Using Magnetic Beads," Science 258:1122-1126 (1992)). Covalent chemical attachment of DNA to the support can be accomplished by using standard coupling agents to link the 5'-phosphate on the DNA to coated microspheres through a phosphoamidate bond. Methods for immobilization of oligonucleotides to solid-state substrates are well established. See Pease *et al.*, Proc. Natl. Acad. Sci. USA 91(11):5022-5026 (1994). One method of attaching oligonucleotides to solid-state substrates is described by Guo *et al.*, Nucleic Acids Res. 22:5456-5465 (1994). Immobilization can be accomplished either by *in situ* DNA synthesis (Maskos and Southern, Nucleic Acids Research, 20:1679-1684 (1992) or by covalent attachment of chemically synthesized oligonucleotides (Guo *et al.*, *supra*) in combination with robotic arraying technologies.

Expression products

[00188] The term "expression products" as used herein refers to both nucleic acids, including, for example, mRNA, and polypeptide products produced by transcription and/or translation of a cancer-associated gene.

[00189] The polypeptides may be in the form of a mature protein or may be a pre-, pro- or prepro- protein that can be activated by cleavage of the pre-, pro- or prepro- portion to produce an active mature polypeptide. In such polypeptides, the pre-, pro- or prepro- sequence may be a leader or secretory sequence or may be a sequence that is employed for purification of the mature polypeptide sequence. Such polypeptides are referred to as "cancer-associated polypeptides".

[00190] The term "cancer-associated polypeptides" also includes variants such as fragments, homologs, fusions and mutants. Homologous polypeptides have at least 80% or more (i.e. at least 85, at least 90, at least 91, at least 92, at least 93, at least 94, at least 95, at least 96, at least 97, at least 98, at least 99%) sequence identity with a cancer-associated polypeptide as referred to above, as determined by the Smith-Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 2, BLOSUM matrix of 62. The Smith-Waterman homology search algorithm is taught in Smith and Waterman, Adv. Appl. Math. (1981) 2: 482-489. The variant polypeptides can be naturally or non-naturally glycosylated, i.e., the polypeptide has a glycosylation pattern that differs from the glycosylation pattern found in the corresponding naturally occurring protein.

[00191] Mutants can include amino acid substitutions, additions or deletions. The amino acid substitutions can be conservative amino acid substitutions or substitutions to eliminate non-essential amino acids, such as to alter a glycosylation site, a phosphorylation site or an

acetylation site, or to minimize misfolding by substitution or deletion of one or more cysteine residues that are not necessary for function. Conservative amino acid substitutions are those that preserve the general charge, hydrophobicity/ hydrophilicity, and/or steric bulk of the amino acid substituted. Variants of these products can be designed so as to retain or have enhanced biological activity of a particular region of the protein (e.g., a functional domain and/or, where the polypeptide is a member of a protein family, a region associated with a consensus sequence). Such variants may then be used in methods of detection or treatment. Selection of amino acid alterations for production of variants can be based upon the accessibility (interior vs. exterior) of the amino acid (see, e.g., Go et al, *Int. J. Peptide Protein Res.* (1980) 15:211), the thermostability of the variant polypeptide (see, e.g., Querol et al., *Prot. Eng.* (1996) 9:265), desired glycosylation sites (see, e.g., Olsen and Thomsen, *J. Gen. Microbiol.* (1991) 137:579), desired disulfide bridges (see, e.g., Clarke et al., *Biochemistry* (1993) 32:4322; and Wakarchuk et al., *Protein Eng.* (1994) 7:1379), desired metal binding sites (see, e.g., Toma et al., *Biochemistry* (1991) 30:97, and Haezebrouck et al., *Protein Eng.* (1993) 6:643), and desired substitutions within proline loops (see, e.g., Masul et al., *Appl. Env. Microbiol.* (1994) 60:3579). Cysteine-depleted muteins can be produced as disclosed in USPN 4,959,314.

[00192] Variants also include fragments of the polypeptides disclosed herein, particularly biologically active fragments and/or fragments corresponding to functional domains. Fragments of interest will typically be at least about 8 amino acids (aa) 10 aa, 15 aa, 20 aa, 25 aa, 30 aa, 35 aa, 40 aa, to at least about 45 aa in length, usually at least about 50 aa in length, at least about 75 aa, at least about 100 aa, at least about 125 aa, at least about 150 aa in length, at least about 200 aa, at least about 300 aa, at least about 400 aa and can be as long as 500 aa in length or longer, but will usually not exceed about 1000 aa in length, where the fragment will have a stretch of amino acids that is identical to a polypeptide encoded by a polynucleotide having a sequence of any one of the polynucleotide sequences provided herein, or a homolog thereof. The protein variants described herein are encoded by polynucleotides that are within the scope of the invention. The genetic code can be used to select the appropriate codons to construct the corresponding variants.

[00193] Altered levels of expression of the polypeptides encoded by cancer-associated genes may indicate that the gene and its products play a role in cancers. In some embodiments, a two-fold increase or decrease in the amount of complex formed is indicative of disease. In some embodiments, a 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold or even 100-fold increase or decrease in the amount of complex formed is indicative of disease.

[00194] Cancer-associated polypeptides may be shorter or longer than the wild type amino acid sequences, and the equivalent coding mRNAs may be similarly modified as compared to the wild type mRNA. Thus, included within the definition of cancer-associated polypeptides are portions or fragments of the wild type sequences herein. In addition, as outlined above, the cancer-associated genes may be used to obtain additional coding regions, and thus additional protein sequence, using techniques known in the art.

[00195] In some embodiments, the cancer-associated polypeptides are derivative or variant cancer-associated polypeptides as compared to the wild-type sequence. That is, as outlined more fully below, the derivative cancer-associated polypeptides will contain at least one amino acid substitution, deletion or insertion. The amino acid substitution, insertion or deletion may occur at any residue within the cancer-associated polypeptides.

[00196] Also included are amino acid sequence variants of cancer-associated polypeptides. These variants fall into one or more of three classes: substitutional, insertional or deletional variants. These variants ordinarily are prepared by site-specific mutagenesis of nucleotides in the DNA encoding the cancer associated protein, using cassette or PCR mutagenesis or other techniques well known in the art, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above. However, variant cancer-associated polypeptide fragments having up to about 100-150 residues may be prepared by in vitro synthesis using established techniques. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the cancer-associated polypeptide amino acid sequence. The variants typically exhibit the same qualitative biological activity as the naturally occurring analogue, although variants can also be selected which have modified characteristics as will be more fully outlined below.

[00197] While the site or region for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed cancer-associated polypeptide variants screened for the optimal combination of desired activity. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, for example, M13 primer mutagenesis and LAR mutagenesis. Screening of the mutants is done using assays of cancer-associated protein activities.

[00198] Amino acid substitutions are typically of single residues, though, of course may be of multiple residues; insertions usually will be on the order of from about 1 to 20 amino acids, although considerably larger insertions may be tolerated. Deletions range from about 1 to about 20 residues, although in some cases deletions may be much larger.

[00199] Substitutions, deletions, insertions or any combination thereof may be used to arrive at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. However, larger changes may be tolerated in certain circumstances. When small alterations in the characteristics of the cancer-associated polypeptide are desired, substitutions are generally made in accordance with the following table:

Table 1

<u>Original Residue</u>	<u>Exemplary Substitutions</u>
Ala	Ser
Arg	Lys
Asn	Gln, His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn, Gln
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	Met, Leu, Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp, Phe
Val	Ile, Leu

[00200] Substantial changes in function or immunological identity occur when substitutions are less conservative than those shown in Table 1. For example, substitutions may be made full

length to more significantly affect one or more of the following: the structure of the polypeptide backbone in the area of the alteration (e.g., the alpha-helical or beta-sheet structure); the charge or hydrophobicity of the molecule at the target site; and the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in the polypeptide's properties are those in which (a) a hydrophilic residue, e.g. seryl or threonyl is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g. lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g. glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g. phenylalanine, is substituted for (or by) one not having a side chain, e.g. glycine.

[00201] The variants typically exhibit the same qualitative biological activity and will elicit the same immune response as the naturally-occurring analogue, although variants may also have modified characteristics.

[00202] The cancer-associated polypeptides may be themselves expressed and used in methods of detection and treatment. They may be further modified in order to assist with their use in such methods.

[00203] Covalent modifications of cancer-associated polypeptides may be utilised, for example in screening. One type of covalent modification includes reacting targeted amino acid residues of a cancer-associated polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a cancer-associated polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking cancer-associated polypeptides to a water-insoluble support matrix or surface for use in the method for purifying anti- cancer-associated antibodies or screening assays, as is more fully described below. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

[00204] Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl, threonyl or tyrosyl residues, methylation of the amino groups of lysine, arginine, and histidine side chains (T.E. Creighton, *Proteins: Structure*

and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)), acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

[00205] Another type of covalent modification of the cancer-associated polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence cancer-associated polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence cancer-associated polypeptide.

[00206] Addition of glycosylation sites to cancer-associated polypeptides may be accomplished by altering the amino acid sequence thereof. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence cancer-associated polypeptide (for O-linked glycosylation sites). The cancer-associated amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the cancer-associated polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

[00207] Another means of increasing the number of carbohydrate moieties on the cancer-associated polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, *LA Crit. Rev. Biochem.*, pp. 259-306 (1981).

[00208] Removal of carbohydrate moieties present on the cancer-associated polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., *Arch. Biochem. Biophys.*, 259:52 (1987) and by Edge et al., *Anal. Biochem.*, 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., *Meth. Enzymol.*, 138:350 (1987).

[00209] Another type of covalent modification of cancer-associated comprises linking the cancer-associated polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

[00210] Cancer-associated polypeptides may also be modified in a way to form chimeric molecules comprising a cancer-associated polypeptide fused to another, heterologous

polypeptide or amino acid sequence. In some embodiments, such a chimeric molecule comprises a fusion of a cancer-associated polypeptide with a tag polypeptide that provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl-terminus of the cancer-associated polypeptide, although internal fusions may also be tolerated in some instances. The presence of such epitope-tagged forms of a cancer-associated polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the cancer-associated polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a cancer-associated polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

[00211] Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 (Field et al., *Mol. Cell. Biol.*, 8:2159-2165 (1988)); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto (Evan et al., *Molecular and Cellular Biology*, 5:3610-3616 (1985)); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky et al., *Protein Engineering*, 3(6):547-553 (1990)). Other tag polypeptides include the Flag-peptide (Hopp et al., *BioTechnology*, 6:1204-1210 (1988)); the KT3 epitope peptide (Martin et al., *Science*, 255:192-194 (1992)); tubulin epitope peptide (Skinner et al., *J. Biol. Chem.*, 266:15163-15166 (1991)); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth et al., *Proc. Natl. Acad. Sci. USA*, 87:6393-6397 (1990)).

[00212] Alternatively, other cancer-associated proteins of the cancer-associated protein family, and cancer-associated proteins from other organisms, may be cloned and expressed as outlined below. Thus, probe or degenerate polymerase chain reaction (PCR) primer sequences may be used to find other related cancer-associated proteins from humans or other organisms. As will be appreciated by those in the art, particularly useful probe and/or PCR primer sequences include the unique areas of the cancer-associated nucleic acid sequence. As is generally known in the art, PCR primers may be from about 15 to about 35 or from about 20 to about 30 nucleotides in length, , and may contain inosine as needed. The conditions for the PCR reaction are well known in the art.

[00213] In addition, as is outlined herein, cancer-associated proteins can be made that are longer than those encoded by cancer-associated genes, for example, by the elucidation of additional

sequences, the addition of epitope or purification tags, the addition of other fusion sequences, etc.

[00214] Cancer-associated proteins may also be identified as being encoded by cancer-associated nucleic acids. Thus, cancer-associated proteins are encoded by nucleic acids that will hybridize to the cancer-associated genes listed above, or their complements, as outlined herein.

Expression of cancer associated polypeptides

[00215] Nucleic acids derived from cancer-associated genes encoding cancer-associated proteins may be used to make a variety of expression vectors to express cancer-associated proteins which can then be used in screening assays, as mentioned above. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate into a host genome. Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the cancer-associated protein. The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

[00216] Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice. The transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the cancer-associated protein; for example, transcriptional and translational regulatory nucleic acid sequences from *Bacillus* are preferably used to express the cancer-associated protein in *Bacillus*. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells.

[00217] In general, the transcriptional and translational regulatory sequences may include, but are not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In some embodiments, the regulatory sequences include a promoter and transcriptional start and stop sequences.

[00218] Promoter sequences encode either constitutive or inducible promoters. The promoters may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention.

[00219] In addition, the expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two organisms, for example in mammalian or insect cells for expression and in a prokaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences that flank the expression construct. The integrating vector may be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art.

[00220] In some embodiments, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes, including antibiotic resistance genes are well known in the art and will vary depending on the host cell used.

[00221] The cancer-associated proteins may be produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a cancer-associated protein, under the appropriate conditions to induce or cause expression of the cancer-associated protein. The conditions appropriate for cancer-associated protein expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

[00222] Appropriate host cells include yeast, bacteria, archaeobacteria, fungi, and insect, plant and animal cells, including mammalian cells. Of particular interest are *Drosophila melanogaster* cells, *Saccharomyces cerevisiae* and other yeasts, *E. coli*, *Bacillus subtilis*, Sf9 cells, C129 cells, 293 cells, *Neurospora*, BHK, CHO, COS, HeLa cells, THP1 cell line (a macrophage cell line) and human cells and cell lines.

[00223] In some embodiments cancer-associated proteins are expressed in mammalian cells. Mammalian expression systems are also known in the art, and include retroviral systems. A preferred expression vector system is a retroviral vector system such as is generally described in WO97/27212 (PCT/US97/01019) and WO97/27213 (PCT/US97/01048), both of which are hereby expressly incorporated by reference. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, herpes simplex virus promoter, and the CMV promoter. Typically, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. Examples of transcription terminator and polyadenylation signals include those derived from SV40.

[00224] The methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, are well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

[00225] In some embodiments, cancer-associated proteins are expressed in bacterial systems. Bacterial expression systems are well known in the art. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters and hybrid promoters are also useful; for example, the tac promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. In addition to a functioning promoter sequence, an efficient ribosome binding site is desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the cancer-associated protein in bacteria. The protein is either secreted into the growth media (Gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (Gram-negative bacteria). The bacterial expression vector may also include a selectable

marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes that render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways. These components are assembled into expression vectors. Expression vectors for bacteria are well known in the art, and include vectors for *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, and *Streptococcus lividans*, among others. The bacterial expression vectors are transformed into bacterial host cells using techniques well known in the art, such as calcium chloride treatment, electroporation, and others.

[00226] Cancer-associated proteins may be produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art.

[00227] In some embodiments, cancer-associated proteins may be produced in yeast cells. Yeast expression systems are well known in the art, and include expression vectors for *Saccharomyces cerevisiae*, *Candida albicans* and *C. maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis* and *K. lactis*, *Pichia guillerimondii* and *P. pastoris*, *Schizosaccharomyces pombe*, and *Yarrowia lipolytica*.

[00228] The cancer-associated protein may also be made as a fusion protein, using techniques well known in the art. Thus, for example, for the creation of monoclonal antibodies. If the desired epitope is small, the cancer-associated protein may be fused to a carrier protein to form an immunogen. Alternatively, the cancer-associated protein may be made as a fusion protein to increase expression, or for other reasons. For example, when the cancer-associated protein is a cancer-associated peptide, the nucleic acid encoding the peptide may be linked to other nucleic acid for expression purposes.

Cancer

[00229] In some embodiments, a cancer detected, diagnosed or treated by the methods of the invention is selected from lymphoma, leukemia, melanoma, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, bladder cancer, stomach cancer or skin cancer.

Antibodies

[00230] In some embodiments the invention uses antibodies that specifically bind to cancer-associated polypeptides expressed by the cancer-associated genes. The term "specifically binds"

means that the antibodies have substantially greater affinity for their target cancer-associated polypeptide than their affinity for other related polypeptides. As used herein, the term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, F(ab')₂ and Fv, which are capable of binding to the antigenic determinant in question. By "substantially greater affinity" we mean that there is a measurable increase in the affinity for the target cancer-associated polypeptide of the invention as compared with the affinity for other related polypeptide. In some embodiments, the affinity is at least 1.5-fold, 2-fold, 5-fold 10-fold, 100-fold, 10³-fold, 10⁴-fold, 10⁵-fold, 10⁶-fold or greater for the target cancer-associated polypeptide.

[00231] In some embodiments, the antibodies bind with high affinity with a dissociation constant of 10⁻⁴M or less, 10⁻⁷M or less, 10⁻⁹M or less or with subnanomolar affinity (0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1 nM or even less).

[00232] When the cancer-associated polypeptides are to be used to generate antibodies, for example for immunotherapy, in some embodiments the cancer-associated polypeptide should share at least one epitope or determinant with the full-length protein. By "epitope" or "determinant" herein is meant a portion of a protein that will generate and/or bind an antibody or T-cell receptor in the context of MHC. Thus, in some instances, antibodies made to a smaller cancer-associated polypeptide will be able to bind to the full-length protein. In some embodiments, the epitope is unique; that is, antibodies generated to a unique epitope show little or no cross-reactivity.

[00233] Polypeptide sequence encoded by the cancer-associated genes may be analyzed to determine certain preferred regions of the polypeptide. Regions of high antigenicity are determined from data by DNASTAR analysis by choosing values that represent regions of the polypeptide that are likely to be exposed on the surface of the polypeptide in an environment in which antigen recognition may occur in the process of initiation of an immune response. For example, the amino acid sequence of a polypeptide encoded by a cancer-associated gene sequence may be analyzed using the default parameters of the DNASTAR computer algorithm (DNASTAR, Inc., Madison, Wis.; see the worldwideweb site at dnastar.com).

[00234] In some embodiments, the antibodies of the present invention bind to orthologs, homologs, paralogs or variants, or combinations and subcombinations thereof, of cancer-associated polypeptides. In some embodiments, the antibodies of the present invention bind to orthologs of cancer-associated polypeptides. In some embodiments, the antibodies of the present invention bind to homologs of cancer-associated polypeptides. In some embodiments, the

antibodies of the present invention bind to paralogs of cancer-associated polypeptides. In some embodiments, the antibodies of the present invention bind to variants of cancer-associated polypeptides. In some embodiments, the antibodies of the present invention do not bind to orthologs, homologs, paralogs or variants, or combinations and subcombinations thereof, of cancer-associated polypeptides.

[00235] Polypeptide features that may be routinely obtained using the DNASTAR computer algorithm include, but are not limited to, Garnier-Robson alpha-regions, beta-regions, turn-regions, and coil-regions (Garnier et al. *J. Mol. Biol.*, 120: 97 (1978)); Chou-Fasman alpha-regions, beta-regions, and turn-regions (*Adv. in Enzymol.*, 47:45-148 (1978)); Kyte-Doolittle hydrophilic regions and hydrophobic regions (*J. Mol. Biol.*, 157:105-132 (1982)); Eisenberg alpha- and beta-amphipathic regions; Karplus-Schulz flexible regions; Emini surface-forming regions (*J. Virol.*, 55(3):836-839 (1985)); and Jameson-Wolf regions of high antigenic index (*CABIOS*, 4(1):181-186 (1988)). Kyte-Doolittle hydrophilic regions and hydrophobic regions, Emini surface-forming regions, and Jameson-Wolf regions of high antigenic index (i.e., containing four or more contiguous amino acids having an antigenic index of greater than or equal to 1.5, as identified using the default parameters of the Jameson-Wolf program) can routinely be used to determine polypeptide regions that exhibit a high degree of potential for antigenicity. One approach for preparing antibodies to a protein is the selection and preparation of an amino acid sequence of all or part of the protein, chemically synthesizing the sequence and injecting it into an appropriate animal, typically a rabbit, hamster or a mouse. Oligopeptides can be selected as candidates for the production of an antibody to the cancer-associated protein based upon the oligopeptides lying in hydrophilic regions, which are thus likely to be exposed in the mature protein. Additional oligopeptides can be determined using, for example, the Antigenicity Index, Welling, G.W. et al., *FEBS Lett.* 188:215-218 (1985), incorporated herein by reference.

[00236] The term "antibody" as used herein includes antibody fragments, as are known in the art, including Fab, Fab2, single chain antibodies (Fv for example), chimeric antibodies, etc., either produced by the modification of whole antibodies or those synthesized de novo using recombinant DNA technologies.

[00237] The invention also provides antibodies that are SMIPs or binding domain immunoglobulin fusion proteins specific for target protein. These constructs are single-chain polypeptides comprising antigen binding domains fused to immunoglobulin domains necessary to carry out antibody effector functions. See e.g., WO03/041600, U.S. Patent publication 20030133939 and US Patent Publication 20030118592.

[00238] Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include a protein encoded by a nucleic acid of the figures or fragment thereof or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants that may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

[00239] In some embodiments the antibodies are monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized in vitro. The immunizing agent will typically include a cancer-associated polypeptide, or fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, *Monoclonal Antibodies: Principles and Practice*, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

[00240] Monoclonal antibody technology is used in implementing research, diagnosis and therapy. Monoclonal antibodies are used in radioimmunoassays, enzyme-linked immunosorbent

assays, immunocytopathology, and flow cytometry for in vitro diagnosis, and in vivo for diagnosis and immunotherapy of human disease. Waldmann, T. A. (1991) *Science* 252:1657-1662. In particular, monoclonal antibodies have been widely applied to the diagnosis and therapy of cancer, wherein it is desirable to target malignant lesions while avoiding normal tissue. See, e.g., U.S. Pat. Nos. 4,753,894 to Frankel, et al.; 4,938,948 to Ring et al.; and 4,956,453 to Bjorn et al.

[00241] The antibodies may be bispecific antibodies. In some embodiments, one of the binding specificities is for a cancer-associated polypeptide, or a fragment thereof, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit, preferably one that is tumor specific.

[00242] In some embodiments, the antibodies to cancer-associated polypeptides are capable of reducing or eliminating the biological function of cancer-associated polypeptides, as is described below. That is, the addition of anti-cancer-associated polypeptide antibodies (either polyclonal or preferably monoclonal) to cancer-associated polypeptides (or cells containing cancer-associated polypeptides) may reduce or eliminate the cancer-associated polypeptide activity. In some embodiments the antibodies of the present invention cause a decrease in activity of at least 25%, at least about 50%, or at least about 95-100%.

[00243] In some embodiments the antibodies to the cancer-associated polypeptides are humanized antibodies. "Humanized" antibodies refer to a molecule having an antigen binding site that is substantially derived from an immunoglobulin from a non-human species and the remaining immunoglobulin structure of the molecule based upon the structure and/or sequence of a human immunoglobulin. The antigen binding site may comprise either complete variable domains fused onto constant domains or only the complementarity determining regions (CDRs) grafted onto appropriate framework regions in the variable domains. Antigen binding sites may be wild type or modified by one or more amino acid substitutions, e.g., modified to resemble human immunoglobulin more closely. Alternatively, a humanized antibody may be derived from a chimeric antibody that retains or substantially retains the antigen-binding properties of the parental, non-human, antibody but which exhibits diminished immunogenicity as compared to the parental antibody when administered to humans. The phrase "chimeric antibody," as used herein, refers to an antibody containing sequence derived from two different antibodies (see, e.g., U.S. Patent No. 4,816,567) that typically originate from different species. Typically, in these chimeric antibodies, the variable region of both light and heavy chains mimics the variable regions of antibodies derived from one species of mammals, while the constant portions are

homologous to the sequences in antibodies derived from another. Most typically, chimeric antibodies comprise human and murine antibody fragments, generally human constant and mouse variable regions. Humanized antibodies are made by replacing the complementarity determining regions (CDRs) of a human antibody (acceptor antibody) with those from a non-human antibody (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human "acceptor" antibody are replaced by corresponding non-human residues from the "donor" antibody. Humanized antibodies may also comprise residues that are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework residues (FR) regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)). One clear advantage to such chimeric forms is that, for example, the variable regions can conveniently be derived from presently known sources using readily available hybridomas or B cells from non human host organisms in combination with constant regions derived from, for example, human cell preparations. While the variable region has the advantage of ease of preparation, and the specificity is not affected by its source, the constant region being human, is less likely to elicit an immune response from a human subject when the antibodies are injected than would the constant region from a non-human source. However, the definition is not limited to this particular example.

[00244] Because humanized antibodies are far less immunogenic in humans than the parental mouse monoclonal antibodies, they can be used for the treatment of humans with far less risk of anaphylaxis. Thus, these antibodies may be preferred in therapeutic applications that involve in vivo administration to a human such as, e.g., use as radiation sensitizers for the treatment of neoplastic disease or use in methods to reduce the side effects of, e.g., cancer therapy. Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source that is non-human. These non-human amino acid residues are often referred to as import residues, which are typically taken from an import variable domain. Humanization can be essentially performed following the method of Winter and co-workers (Jones et al., *Nature* 321:522-525 (1986);

Riechmann et al., *Nature* 332:323-327 (1988); Verhoeyen et al., *Science* 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

[00245] A number of "humanized" antibody molecules comprising an antigen-binding site derived from a non-human immunoglobulin have been described, including chimeric antibodies having rodent V regions and their associated CDRs fused to human constant domains (Winter et al. (1991) *Nature* 349:293-299; Lobuglio et al. (1989) *Proc. Nat. Acad. Sci. USA* 86:4220-4224; Shaw et al. (1987) *J Immunol.* 138:4534-4538; and Brown et al. (1987) *Cancer Res.* 47:3577-3583), rodent CDRs grafted into a human supporting FR prior to fusion with an appropriate human antibody constant domain (Riechmann et al. (1988) *Nature* 332:323-327; Verhoeyen et al. (1988) *Science* 239:1534-1536; and Jones et al. (1986) *Nature* 321:522-525), and rodent CDRs supported by recombinantly veneered rodent FRs (European Patent Publication No. 519,596, published Dec. 23, 1992).

[00246] Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); Marks et al., *J. Mol. Biol.*, 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77 (1985) and Boerner et al., *J. Immunol.*, 147(1):86-95 (1991)). Humanized antibodies may be achieved by a variety of methods including, for example: (1) grafting the non-human complementarity determining regions (CDRs) onto a human framework and constant region (a process referred to in the art as "humanizing"), or, alternatively, (2) transplanting the entire non-human variable domains, but "cloaking" them with a human-like surface by replacement of surface residues (a process referred to in the art as "veneering"). In the present invention, humanized antibodies will include both "humanized" and "veneered" antibodies. Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for

example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks et al., *Bio/Technology* 10, 779-783 (1992); Lonberg et al., *Nature* 368 856-859 (1994); Morrison, *Nature* 368, 812-13 (1994); Fishwild et al., *Nature Biotechnology* 14, 845-51 (1996); Neuberger, *Nature Biotechnology* 14, 826 (1996); Lonberg and Huszar, *Intern. Rev. Immunol.* 13 65-93 (1995); Jones et al., *Nature* 321:522-525 (1986); Morrison et al., *Proc. Natl. Acad. Sci. US.A.*, 81:6851-6855 (1984); Morrison and Oi, *Adv. Immunol.*, 44:65-92 (1988); Verhoeyer et al., *Science* 239:1534-1536 (1988); Padlan, *Molec. Immun.* 28:489-498 (1991); Padlan, *Molec. Immunol.* 31(3):169-217 (1994); and Kettleborough, C.A. et al., *Protein Eng.* 4(7):773-83 (1991) each of which is incorporated herein by reference. Antibodies of the present invention can also be produced using human engineering techniques as discussed in U.S. Patent 5,766,886, which is incorporated herein by reference.

[00247] The phrase "complementarity determining region" refers to amino acid sequences which together define the binding affinity and specificity of the natural Fv region of a native immunoglobulin binding site. See, e.g., Chothia et al., *J. Mol. Biol.* 196:901-917 (1987); Kabat et al., U.S. Dept. of Health and Human Services NIH Publication No. 91-3242 (1991). The phrase "constant region" refers to the portion of the antibody molecule that confers effector functions. In the present invention, mouse constant regions are substituted by human constant regions. The constant regions of the subject humanized antibodies are derived from human immunoglobulins. The heavy chain constant region can be selected from any of the five isotypes: alpha, delta, epsilon, gamma or mu. One method of humanizing antibodies comprises aligning the non-human heavy and light chain sequences to human heavy and light chain sequences, selecting and replacing the non-human framework with a human framework based on such alignment, molecular modeling to predict the conformation of the humanized sequence and comparing to the conformation of the parent antibody. This process is followed by repeated back mutation of residues in the CDR region that disturb the structure of the CDRs until the predicted conformation of the humanized sequence model closely approximates the conformation of the non-human CDRs of the parent non-human antibody. Such humanized antibodies may be further derivatized to facilitate uptake and clearance, e.g. via Ashwell receptors. See, e.g., U.S. Patent Nos. 5,530,101 and 5,585,089 which are incorporated herein by reference.

[00248] Humanized antibodies to cancer-associated polypeptides can also be produced using transgenic animals that are engineered to contain human immunoglobulin loci. For example, WO 98/24893 discloses transgenic animals having a human Ig locus wherein the animals do not produce functional endogenous immunoglobulins due to the inactivation of endogenous heavy

and light chain loci. WO 91/10741 also discloses transgenic non-primate mammalian hosts capable of mounting an immune response to an immunogen, wherein the antibodies have primate constant and/or variable regions, and wherein the endogenous immunoglobulin-encoding loci are substituted or inactivated. WO 96/30498 discloses the use of the Cre/Lox system to modify the immunoglobulin locus in a mammal, such as to replace all or a portion of the constant or variable region to form a modified antibody molecule. WO 94/02602 discloses non-human mammalian hosts having inactivated endogenous Ig loci and functional human Ig loci. U.S. Patent No. 5,939,598 discloses methods of making transgenic mice in which the mice lack endogenous heavy chains, and express an exogenous immunoglobulin locus comprising one or more xenogeneic constant regions.

[00249] Using a transgenic animal described above, an immune response can be produced to a selected antigenic molecule, and antibody-producing cells can be removed from the animal and used to produce hybridomas that secrete human monoclonal antibodies. Immunization protocols, adjuvants, and the like are known in the art, and are used in immunization of, for example, a transgenic mouse as described in WO 96/33735. The monoclonal antibodies can be tested for the ability to inhibit or neutralize the biological activity or physiological effect of the corresponding protein.

[00250] In some embodiments, cancer-associated polypeptides as recited above and variants thereof may be used to immunize a transgenic animal as described above. Monoclonal antibodies are made using methods known in the art, and the specificity of the antibodies is tested using isolated cancer-associated polypeptides. Methods for preparation of the human or primate cancer-associated or an epitope thereof include, but are not limited to chemical synthesis, recombinant DNA techniques or isolation from biological samples. Chemical synthesis of a peptide can be performed, for example, by the classical Merrifield method of solid phase peptide synthesis (Merrifield, J. Am. Chem. Soc. 85:2149, 1963 which is incorporated by reference) or the FMOC strategy on a Rapid Automated Multiple Peptide Synthesis system (E. I. du Pont de Nemours Company, Wilmington, DE) (Caprino and Han, J. Org. Chem. 37:3404, 1972 which is incorporated by reference).

[00251] Polyclonal antibodies can be prepared by immunizing rabbits or other animals by injecting antigen followed by subsequent boosts at appropriate intervals. Alternative animals include mice, rats, chickens, guinea pigs, sheep, horses, monkeys, camels and sharks. The animals are bled and sera assayed against purified cancer-associated proteins usually by ELISA or by bioassay based upon the ability to block the action of cancer-associated proteins. When

using avian species, e.g., chicken, turkey and the like, the antibody can be isolated from the yolk of the egg. Monoclonal antibodies can be prepared after the method of Milstein and Kohler by fusing splenocytes from immunized mice with continuously replicating tumor cells such as myeloma or lymphoma cells. (Milstein and Kohler, *Nature* 256:495-497, 1975; Galfre and Milstein, *Methods in Enzymology: Immunochemical Techniques* 73:1-46, Langone and Banatis eds., Academic Press, 1981 which are incorporated by reference). The hybridoma cells so formed are then cloned by limiting dilution methods and supernates assayed for antibody production by ELISA, RIA or bioassay.

[00252] The unique ability of antibodies to recognize and specifically bind to target proteins provides an approach for treating an overexpression of the protein. Thus, in some embodiments the present invention provides methods for preventing or treating diseases involving overexpression of a cancer-associated polypeptide by treatment of a patient with specific antibodies to the cancer-associated protein.

[00253] Specific antibodies, either polyclonal or monoclonal, to the cancer-associated proteins can be produced by any suitable method known in the art as discussed above. For example, murine or human monoclonal antibodies can be produced by hybridoma technology or, alternatively, the cancer-associated proteins, or an immunologically active fragment thereof, or an anti-idiotypic antibody, or fragment thereof can be administered to an animal to elicit the production of antibodies capable of recognizing and binding to the cancer-associated proteins. Such antibodies can be from any class of antibodies including, but not limited to IgG, IgA, IgM, IgD, and IgE or in the case of avian species, IgY and from any subclass of antibodies.

[00254] In some embodiments the antibodies of the present invention are neutralizing antibodies. In some embodiments the antibodies are targeting antibodies. In some embodiments, the antibodies are internalized upon binding a target. In some embodiments the antibodies do not become internalized upon binding a target and instead remain on the surface.

[00255] The antibodies of the present invention can be screened for the ability to either be rapidly internalized upon binding to the tumor-cell antigen in question, or for the ability to remain on the cell surface following binding. In some embodiments, for example in the construction of some types of immunoconjugates, the ability of an antibody to be internalized may be desired if internalization is required to release the toxin moiety. Alternatively, if the antibody is being used to promote ADCC or CDC, it may be more desirable for the antibody to remain on the cell surface. A screening method can be used to differentiate these type behaviors.

For example, a tumor cell antigen bearing cell may be used where the cells are incubated with human IgG1 (control antibody) or one of the antibodies of the invention at a concentration of approximately 1 $\mu\text{g/mL}$ on ice (with 0.1% sodium azide to block internalization) or 37°C (without sodium azide) for 3 hours. The cells are then washed with cold staining buffer (PBS+1%BSA+0.1% sodium azide), and are stained with goat anti-human IgG-FITC for 30 minutes on ice. Geometric mean fluorescent intensity (MFI) is recorded by FACS Calibur. If no difference in MFI is observed between cells incubated with the antibody of the invention on ice in the presence of sodium azide and cells observed at 37°C in the absence of sodium azide, the antibody will be suspected to be one that remains bound to the cell surface, rather than being internalized. If however, a decrease in surface stainable antibody is found when the cells are incubated at 37°C in the absence of sodium azide, the antibody will be suspected to be one which is capable of internalization.

Antibody Conjugates

[00256] In some embodiments, the antibodies of the invention are conjugated. In some embodiments, the conjugated antibodies are useful for cancer therapeutics, cancer diagnosis, or imaging of cancerous cells.

[00257] For diagnostic applications, the antibody typically will be labeled with a detectable moiety. Numerous labels are available which can be generally grouped into the following categories:

(a) Radionuclides such as those discussed *infra*. The antibody can be labeled, for example, with the radioisotope using the techniques described in Current Protocols in Immunology, Volumes 1 and 2, Coligan et al., Ed. Wiley-Interscience, New York, N.Y., Pubs. (1991) for example and radioactivity can be measured using scintillation counting.

(b) Fluorescent labels such as rare earth chelates (europium chelates) or fluorescein and its derivatives, rhodamine and its derivatives, dansyl, Lissamine, phycoerythrin and Texas Red are available. The fluorescent labels can be conjugated to the antibody using the techniques disclosed in Current Protocols in Immunology, *supra*, for example. Fluorescence can be quantified using a fluorimeter.

(c) Various enzyme-substrate labels are available and U.S. Pat. No. 4,275,149 provides a review of some of these. The enzyme generally catalyzes a chemical alteration of the chromogenic substrate which can be measured using various techniques. For example, the enzyme may catalyze a color change in a substrate, which can be measured

spectrophotometrically. Alternatively, the enzyme may alter the fluorescence or chemiluminescence of the substrate. Techniques for quantifying a change in fluorescence are described above. The chemiluminescent substrate becomes electronically excited by a chemical reaction and may then emit light which can be measured (using a chemiluminometer, for example) or donates energy to a fluorescent acceptor. Examples of enzymatic labels include luciferases (e.g., firefly luciferase and bacterial luciferase; U.S. Pat. No. 4,737,456), luciferin, 2,3-dihydrophthalazinediones, malate dehydrogenase, urease, peroxidase such as horseradish peroxidase (HRPO), alkaline phosphatase, .beta.-galactosidase, glucoamylase, lysozyme, saccharide oxidases (e.g., glucose oxidase, galactose oxidase, and glucose-6-phosphate dehydrogenase), heterocyclic oxidases (such as uricase and xanthine oxidase), lactoperoxidase, microperoxidase, and the like. Techniques for conjugating enzymes to antibodies are described in O'Sullivan et al., *Methods for the Preparation of Enzyme-Antibody Conjugates for use in Enzyme Immunoassay*, in *Methods in Enzym.* (ed J. Langone & H. Van Vunakis), Academic press, New York, 73:147-166 (1981).

[00258] The antibodies may also be used for in vivo diagnostic assays. In some embodiments, the antibody is labeled with a radionuclide so that the tumor can be localized using immunoscintigraphy. As a matter of convenience, the antibodies of the present invention can be provided in a kit, i. e., a packaged combination of reagents in predetermined amounts with instructions for performing the diagnostic assay. Where the antibody is labeled with an enzyme, the kit may include substrates and cofactors required by the enzyme (e.g., a substrate precursor which provides the detectable chromophore or fluorophore). In addition, other additives may be included such as stabilizers, buffers (e.g., a block buffer or lysis buffer) and the like. The relative amounts of the various reagents may be varied widely to provide for concentrations in solution of the reagents which substantially optimize the sensitivity of the assay. Particularly, the reagents may be provided as dry powders, usually lyophilized, including excipients which on dissolution will provide a reagent solution having the appropriate concentration.

[00259] In some embodiments, antibodies are conjugated to one or more maytansine molecules (e.g. about 1 to about 10 maytansine molecules per antibody molecule). Maytansine may, for example, be converted to May-SS-Me which may be reduced to May-SH₃ and reacted with modified antibody (Chari et al. *Cancer Research* 52: 127-131 (1992)) to generate a maytansinoid-antibody immunoconjugate. In some embodiments, the conjugate may be the highly potent maytansine derivative DM1 (N²'-deacetyl-N²'-(3-mercapto-1-oxopropyl)-maytansine) (see for example WO02/098883 published Dec. 12, 2002) which has an IC₅₀ of

approximately 10-11 M (review, see Payne (2003) *Cancer Cell* 3:207- 212) or DM4 (N2'-deacetyl-N2'(4-methyl- -4-mercapto-1-oxopentyl)-maytansine) (see for example WO2004/103272 published Dec. 2, 2004).

[00260] In some embodiments the antibody conjugate comprises an anti-tumor cell antigen antibody conjugated to one or more calicheamicin molecules. The calicheamicin family of antibiotics are capable of producing double-stranded DNA breaks at sub-picomolar concentrations. Structural analogues of calicheamicin which may be used include, but are not limited to, gamma₁I, alpha₂I, alpha₃I, N-acetyl-gamma₁I, PSAG and thetaI₁ (Hinman et al. *Cancer Research* 53: 3336-3342 (1993) and Lode et al. *Cancer Research* 58: 2925-2928 (1998)). See, also, U.S. Pat. Nos. 5,714,586; 5,712,374; 5,264,586; and 5,773,001, each of which is expressly incorporated herein by reference.

[00261] In some embodiments the antibody is conjugated to a prodrug capable of being release in its active form by enzymes overproduced in many cancers. For example, antibody conjugates can be made with a prodrug form of doxorubicin wherein the active component is released from the conjugate by plasmin. Plasmin is known to be over produced in many cancerous tissues (see Decy et al, (2004) *FASEB Journal* 18(3): 565-567).

[00262] In some embodiments the antibodies are conjugated to enzymatically active toxins and fragments thereof. In some embodiments the toxins include, without limitation, diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), *Pseudomonas* endotoxin, ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), Ribonuclease (Rnase), Deoxyribonuclease (Dnase), pokeweed antiviral protein, momordica charantia inhibitor, curcin, croton, *sapaonaria officinalis* inhibitor, gelonin, mitogellin, restrictocin, phenomycin, neomycin and the tricothecenes. See, for example, WO 93/21232 published Oct. 28, 1993. In some embodiments the toxins have low intrinsic immunogenicity and a mechanism of action (e.g. a cytotoxic mechanism versus a cytostatic mechanism) that reduces the opportunity for the cancerous cells to become resistant to the toxin.

[00263] In some embodiments conjugates made between the antibodies of the invention and immunomodulators. For example, in some embodiments immunostimulatory oligonucleotides can be used. These molecules are potent immunogens that can elicit antigen-specific antibody responses (see Datta et al, (2003) *Ann N.Y. Acad. Sci* 1002: 105-111). Additional immunomodulatory compounds can include stem cell growth factor such as "S1 factor",

lymphotoxins such as tumor necrosis factor (TNF), hematopoietic factor such as an interleukin, colony stimulating factor (CSF) such as granulocyte-colony stimulating factor (G-CSF) or granulocyte macrophage-stimulating factor (GM-CSF), interferon (IFN) such as interferon alpha, beta or gamma, erythropoietin, and thrombopoietin.

[00264] In some embodiments radioconjugated antibodies are provided. In some embodiments such antibodies can be made using P-32, P-33, Sc-47, Fe-59, Cu-64, Cu-67, Se-75, As-77, Sr-89, Y-90, Mo-99, Rh-105, Pd-109, Ag-111, I-125, I-131, Pr-142, Pr-143, Pm-149, Sm-153, Th-161, Ho-166, Er-169, Lu-177, Re-186, Re-188, Re-189, Ir-194, Au-198, Au-199, Pb-211, Pb-212, and Bi-213, Co-58, Ga-67, Br-80m, Tc-99m, Rh-103m, Pt-109, In-111, Sb-119, I-125, Ho-161, Os-189m, Ir-192, Dy-152, At-211, Bi-212, Ra-223, Rn-219, Po-215, Bi-211, Ac-225, Fr-221, At-217, Bi-213, Fm-255 and combinations and subcombinations thereof. In some embodiments, boron, gadolinium or uranium atoms are conjugated to the antibodies. In some embodiments the boron atom is B-10, the gadolinium atom is Gd-157 and the uranium atom is U-235.

[00265] In some embodiments the radionuclide conjugate has a radionuclide with an energy between 20 and 10,000 keV. The radionuclide can be an Auger emitter, with an energy of less than 1000 keV, a P emitter with an energy between 20 and 5000 keV, or an alpha or 'a' emitter with an energy between 2000 and 10,000 keV.

[00266] In some embodiments diagnostic radioconjugates are provided which comprise a radionuclide that is a gamma- beta- or positron-emitting isotope. In some embodiments the radionuclide has an energy between 20 and 10,000 keV. In some embodiments the radionuclide is selected from the group of ^{18}F , ^{51}Mn , $^{52\text{m}}\text{Mn}$, ^{52}Fe , ^{55}Co , ^{62}Cu , ^{64}Cu , ^{68}Ga , ^{72}As , ^{75}Br , ^{76}Br , $^{82\text{m}}\text{Rb}$, ^{83}Sr , ^{86}Y , ^{89}Zr , $^{94\text{m}}\text{Tc}$, ^{51}Cr , ^{57}Co , ^{58}Co , ^{59}Fe , ^{67}Cu , ^{67}Ga , ^{75}Se , ^{97}Ru , $^{99\text{m}}\text{Tc}$, $^{114\text{m}}\text{In}$, ^{123}I , ^{125}I , ^{13}Li and ^{197}Hg .

[00267] In some embodiments the antibodies of the invention are conjugated to diagnostic agents that are photoactive or contrast agents. Photoactive compounds can comprise compounds such as chromagens or dyes. Contrast agents may be, for example a paramagnetic ion, wherein the ion comprises a metal selected from the group of chromium (III), manganese (II), iron (III), iron (II), cobalt (II), nickel (II), copper (II), neodymium (III), samarium (III), ytterbium (III), gadolinium (III), vanadium (II), terbium (III), dysprosium (III), holmium (III) and erbium (III). The contrast agent may also be a radio-opaque compound used in X-ray techniques or computed tomography, such as an iodine, iridium, barium, gallium and thallium compound. Radio-opaque

compounds may be selected from the group of barium, diatrizoate, ethiodized oil, gallium citrate, iocarmic acid, iocetamic acid, iodamide, iodipamide, iodoxamic acid, iogulamide, iohexol, iopamidol, iopanoic acid, ioprocemic acid, iosefamic acid, ioseric acid, iosulamide meglumine, iosemetic acid, iotasul, iotetric acid, iothalamic acid, iotroxic acid, ioxaglic acid, ioxotrizoic acid, ipodate, meglumine, metrizamide, metrizoate, propyliodone, and thallous chloride. In some embodiments, the diagnostic immunoconjugates may contain ultrasound-enhancing agents such as a gas filled liposome that is conjugated to an antibody of the invention. Diagnostic immunoconjugates may be used for a variety of procedures including, but not limited to, intraoperative, endoscopic or intravascular methods of tumor or cancer diagnosis and detection.

[00268] In some embodiments antibody conjugates are made using a variety of bifunctional protein coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate, iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al. Science 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026. The linker may be a "cleavable linker" facilitating release of the cytotoxic drug in the cell. For example, an acid-labile linker, peptidase-sensitive linker, dimethyl linker or disulfide-containing linker (Chari et al. Cancer Research 52: 127-131 (1992)) may be used. Agents may be additionally be linked to the antibodies of the invention through a carbohydrate moiety.

[00269] In some embodiments fusion proteins comprising the antibodies of the invention and cytotoxic agents may be made, e.g. by recombinant techniques or peptide synthesis. In some embodiments such immunoconjugates comprising the anti-tumor antigen antibody conjugated with a cytotoxic agent are administered to the patient. In some embodiments the immunoconjugate and/or tumor cell antigen protein to which it is bound is/are internalized by the cell, resulting in increased therapeutic efficacy of the immunoconjugate in killing the cancer cell to which it binds. In some embodiments, the cytotoxic agent targets or interferes with nucleic

acid in the cancer cell. Examples of such cytotoxic agents include maytansinoids, calicheamicins, ribonucleases and DNA endonucleases.

[00270] In some embodiments the antibodies are conjugated to a "receptor" (such as streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g. avidin) which is conjugated to a cytotoxic agent (e.g. a radionucleotide).

[00271] In some embodiments the antibodies are conjugated to a cytotoxic molecule which is released inside a target cell lysosome. For example, the drug monomethyl auristatin E (MMAE) can be conjugated via a valine-citrulline linkage which will be cleaved by the proteolytic lysosomal enzyme cathepsin B following internalization of the antibody conjugate (see for example WO03/026577 published April 3, 2003). In some embodiments, the MMAE can be attached to the antibody using an acid-labile linker containing a hydrazone functionality as the cleavable moiety (see for example WO02/088172 published Nov. 11, 2002).

Antibody Dependent Enzyme Mediated Prodrug Therapy (ADEPT)

[00272] In some embodiments the antibodies of the present invention may also be used in ADEPT by conjugating the antibody to a prodrug-activating enzyme which converts a prodrug (e.g. a peptidyl chemotherapeutic agent, see WO81/01145) to an active anti-cancer drug. See, for example, WO 88/07378 and U.S. Pat. No. 4,975,278.

[00273] In some embodiments the enzyme component of the immunoconjugate useful for ADEPT includes any enzyme capable of acting on a prodrug in such a way so as to convert it into its more active, cytotoxic form.

[00274] Enzymes that are useful in ADEPT include, but are not limited to, alkaline phosphatase useful for converting phosphate-containing prodrugs into free drugs; arylsulfatase useful for converting sulfate-containing prodrugs into free drugs; cytosine deaminase useful for converting non-toxic 5-fluorocytosine into the anti-cancer drug, 5-fluorouracil; proteases, such as serratio protease, thermolysin, subtilisin, carboxypeptidases and cathepsins (such as cathepsins B and L), that are useful for converting peptide-containing prodrugs into free drugs; D-alanylcarboxypeptidases, useful for converting prodrugs that contain D-amino acid substituents; carbohydrate-cleaving enzymes such as β -galactosidase and neuraminidase useful for converting glycosylated prodrugs into free drugs; β -lactamase useful for converting drugs derivatized with β -lactams into free drugs; and penicillin amidases, such as penicillin V

amidase or penicillin G amidase, useful for converting drugs derivatized at their amine nitrogens with phenoxyacetyl or phenylacetyl groups, respectively, into free drugs. In some embodiments antibodies with enzymatic activity, also known in the art as "abzymes", can be used to convert the prodrugs of the invention into free active drugs (see, e.g., Massey, *Nature* 328: 457-458 (1987)). Antibody-abzyme conjugates can be prepared as described herein for delivery of the abzyme to a tumor cell population.

[00275] In some embodiments the ADEPT enzymes can be covalently bound to the antibodies by techniques well known in the art such as the use of the heterobifunctional crosslinking reagents discussed above. In some embodiments, fusion proteins comprising at least the antigen binding region of an antibody of the invention linked to at least a functionally active portion of an enzyme of the invention can be constructed using recombinant DNA techniques well known in the art (see, e.g., Neuberger et al., *Nature*, 312: 604-608 (1984)).

[00276] In some embodiments identification of an antibody that acts in a cytostatic manner rather than an cytotoxic manner can be accomplished by measuring viability of a treated target cell culture in comparison with a non-treated control culture. Viability can be detected using methods known in the art such as the CellTiter-Blue® Cell Viability Assay or the CellTiter-Glo® Luminescent Cell Viability Assay (Promega, catalog numbers G8080 and G5750 respectively). In some embodiments an antibody is considered as potentially cytostatic if treatment causes a decrease in cell number in comparison to the control culture without any evidence of cell death as measured by the means described above.

[00277] In some embodiments an in vitro screening assay can be performed to identify an antibody that promotes ADCC using assays known in the art. One exemplary assay is the In Vitro ADCC Assay. To prepare chromium 51-labeled target cells, tumor cell lines are grown in tissue culture plates and harvested using sterile 10 mM EDTA in PBS. The detached cells are washed twice with cell culture medium. Cells (5×10^6) are labeled with 200 μ Ci of chromium 51 (New England Nuclear/DuPont) at 37° C. for one hour with occasional mixing. Labeled cells were washed three times with cell culture medium, then are resuspended to a concentration of 1×10^5 cells/mL. Cells are used either without opsonization, or are opsonized prior to the assay by incubation with test antibody at 100 ng/mL and 1.25 ng/mL in PBMC assay or 20 ng/mL and 1 ng/mL in NK assay. Peripheral blood mononuclear cells are prepared by collecting blood on heparin from normal healthy donors and diluted with an equal volume of phosphate buffered saline (PBS). The blood is then layered over LYMPHOCYTE SEPARATION MEDIUM® (LSM: Organon Teknika) and centrifuged according to the manufacturer's instructions.

Mononuclear cells are collected from the LSM-plasma interface and are washed three times with PBS. Effector cells are suspended in cell culture medium to a final concentration of 1×10^7 cells/mL. After purification through LSM, natural killer (NK) cells are isolated from PBMCs by negative selection using an NK cell isolation kit and a magnetic column (Milenyi Biotech) according to the manufacturer's instructions. Isolated NK cells are collected, washed and resuspended in cell culture medium to a concentration of 2×10^6 cells/mL. The identity of the NK cells is confirmed by flow cytometric analysis. Varying effector:target ratios are prepared by serially diluting the effector (either PBMC or NK) cells two-fold along the rows of a microtiter plate (100 μ L final volume) in cell culture medium. The concentration of effector cells ranges from 1.0×10^7 /mL to 2.0×10^4 /mL for PBMC and from 2.0×10^6 /mL to 3.9×10^3 /mL for NK. After titration of effector cells, 100 μ L of chromium 51-labeled target cells (opsonized or nonopsonized) at 1×10^5 cells/mL are added to each well of the plate. This results in an initial effector:target ratio of 100:1 for PBMC and 20:1 for NK cells. All assays are run in duplicate, and each plate contains controls for both spontaneous lysis (no effector cells) and total lysis (target cells plus 100 μ L 1% sodium dodecyl sulfate, 1 N sodium hydroxide). The plates are incubated at 37° C. for 18 hours, after which the cell culture supernatants are harvested using a supernatant collection system (Skatron Instrument, Inc.) and counted in a Minaxi auto-gamma 5000 series gamma counter (Packard) for one minute. Results are then expressed as percent cytotoxicity using the formula: % Cytotoxicity=(sample cpm-spontaneous lysis)/(total lysis-spontaneous lysis) \times 100.

[00278] To identify an antibody that promotes CDC, the skilled artisan may perform an assay known in the art. One exemplary assay is the In Vitro CDC assay. In vitro, CDC activity can be measured by incubating tumor cell antigen expressing cells with human (or alternate source) complement-containing serum in the absence or presence of different concentrations of test antibody. Cytotoxicity is then measured by quantifying live cells using ALAMAR BLUE® (Gazzano-Santoro et al., J. Immunol. Methods 202 163-171 (1997)). Control assays are performed without antibody, and with antibody, but using heat inactivated serum and/or using cells which do not express the tumor cell antigen in question. Alternatively, red blood cells can be coated with tumor antigen or peptides derived from tumor antigen, and then CDC may be assayed by observing red cell lysis (see for example Karjalainen and Mantylarvi, Acta Pathol Microbiol Scand [C]. 1981 Oct;89(5):315-9).

[00279] To select for antibodies that induce cell death, loss of membrane integrity as indicated by, e.g., PI, trypan blue or 7AAD uptake may be assessed relative to control. One exemplary

assay is the PI uptake assay using tumor antigen expressing cells. According to this assay, tumor cell antigen expressing cells are cultured in Dulbecco's Modified Eagle Medium (D-MEM):Ham's F-12 (50:50) supplemented with 10% heat-inactivated FBS (Hyclone) and 2 mM L-glutamine. (Thus, the assay is performed in the absence of complement and immune effector cells). The tumor cells are seeded at a density of 3×10^6 per dish in 100 x 20 mm dishes and allowed to attach overnight. The medium is then removed and replaced with fresh medium alone or medium containing 10 $\mu\text{g/mL}$ of the appropriate monoclonal antibody. The cells are incubated for a 3 day time period. Following each treatment, monolayers are washed with PBS and detached by trypsinization. Cells are then centrifuged at 1200 rpm for 5 minutes at 4.degree. C., the pellet resuspended in 3 mL ice cold Ca^{2+} binding buffer (10 mM Hepes, pH 7.4, 140 mM NaCl, 2.5 mM CaCl_2) and aliquoted into 35 mm strainer-capped 12 x 75 tubes (1 mL per tube, 3 tubes per treatment group) for removal of cell clumps. Tubes then receive PI (10 $\mu\text{g/mL}$). Samples may be analyzed using a FACSCANTM flow cytometer and FACSCONVERTTM. CellQuest software (Becton Dickinson). Those antibodies that induce statistically significant levels of cell death as determined by PI uptake may be selected as cell death-inducing antibodies.

[00280] Antibodies can also be screened in vivo for apoptotic activity using ^{18}F -annexin as a PET imaging agent. In this procedure, Annexin V is radiolabeled with ^{18}F and given to the test animal following dosage with the antibody under investigation. One of the earliest events to occur in the apoptotic process is the eversion of phosphatidylserine from the inner side of the cell membrane to the outer cell surface, where it is accessible to annexin. The animals are then subjected to PET imaging (see Yagle et al, J Nucl Med. 2005 Apr;46(4):658-66). Animals can also be sacrificed and individual organs or tumors removed and analyzed for apoptotic markers following standard protocols.

[00281] While in some embodiments cancer may be characterized by overexpression of a gene expression product, the present application further provides methods for treating cancer which is not considered to be a tumor antigen-overexpressing cancer. To determine tumor antigen expression in the cancer, various diagnostic/prognostic assays are available. In some embodiments, gene expression product overexpression can be analyzed by IHC. Paraffin embedded tissue sections from a tumor biopsy may be subjected to the IHC assay and accorded a tumor antigen protein staining intensity criteria as follows:

Score 0: no staining is observed or membrane staining is observed in less than 10% of tumor cells.

Score 1+: a faint/barely perceptible membrane staining is detected in more than 10% of the tumor cells. The cells are only stained in part of their membrane.

Score 2+: a weak to moderate complete membrane staining is observed in more than 10% of the tumor cells.

Score 3+: a moderate to strong complete membrane staining is observed in more than 10% of the tumor cells.

[00282] Those tumors with 0 or 1+ scores for tumor antigen overexpression assessment may be characterized as not overexpressing the tumor antigen, whereas those tumors with 2+ or 3+ scores may be characterized as overexpressing the tumor antigen.

[00283] Alternatively, or additionally, FISH assays such as the INFORM™ (sold by Ventana, Ariz.) or PATHVISION™ (Vysis, Ill.) may be carried out on formalin-fixed, paraffin-embedded tumor tissue to determine the extent (if any) of tumor antigen overexpression in the tumor.

[00284] Additionally, antibodies can be chemically modified by covalent conjugation to a polymer to increase their circulating half-life, for example. Each antibody molecule may be attached to one or more (i.e. 1, 2, 3, 4, 5 or more) polymer molecules. Polymer molecules are preferably attached to antibodies by linker molecules. The polymer may, in general, be a synthetic or naturally occurring polymer, for example an optionally substituted straight or branched chain polyalkene, polyalkenylene or polyoxyalkylene polymer or a branched or unbranched polysaccharide, e.g. homo- or hetero-polysaccharide. In some embodiments the polymers are polyoxyethylene polyols and polyethylene glycol (PEG). PEG is soluble in water at room temperature and has the general formula: $R(O-CH_2-CH_2)_nO-R$ where R can be hydrogen, or a protective group such as an alkyl or alkanol group. In some embodiments, the protective group has between 1 and 8 carbons. In some embodiments the protective group is methyl. The symbol n is a positive integer, between 1 and 1,000, or 2 and 500. In some embodiments the PEG has an average molecular weight between 1000 and 40,000, between 2000 and 20,000, or between 3,000 and 12,000. In some embodiments, PEG has at least one hydroxy group. In some embodiments the hydroxy is a terminal hydroxy group. In some embodiments it is this hydroxy group which is activated to react with a free amino group on the inhibitor. However, it will be understood that the type and amount of the reactive groups may be varied to achieve a covalently conjugated PEG/antibody of the present invention. Polymers, and methods to attach them to peptides, are shown in U. S. Patent Numbers 4,766, 106; 4,179, 337; 4,495, 285; and 4,609, 546 each of which is hereby incorporated by reference in its entirety.

Labelling and Detection

[00285] In some embodiments, the cancer-associated nucleic acids, proteins and antibodies of the invention are labeled. It is noted that many of the examples of conjugates discussed *supra* are also relevant to non-antibodies. To the extent such examples are relevant they are incorporated herein.

[00286] By "labeled" herein is meant that a compound has at least one element, isotope or chemical compound attached to enable the detection of the compound. In general, labels fall into three classes: a) isotopic labels, which may be radioactive or heavy isotopes; b) immune labels, which may be antibodies or antigens; and c) coloured or fluorescent dyes. The labels may be incorporated into the cancer-associated nucleic acids, proteins and antibodies at any position. For example, the label should be capable of producing, either directly or indirectly, a detectable signal. The detectable moiety may be a radioisotope, such as ^3H , ^{14}C , ^{32}P , ^{35}S , or ^{125}I , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the label may be employed, including those methods described by Hunter et al., *Nature*, 144:945 (1962); David et al., *Biochemistry*, 13:1014 (1974); Pain et al., *J. Immunol. Meth.*, 40:219 (1981); and Nygren, J. *Histochem. and Cytochem.*, 30:407 (1982).

[00287] Detection of the expression product of interest can be accomplished using any detection method known to those of skill in the art. "Detecting expression" or "detecting the level of" is intended to mean determining the quantity or presence of a biomarker protein or gene in the biological sample. Thus, "detecting expression" encompasses instances where a biomarker is determined not to be expressed, not to be detectably expressed, expressed at a low level, expressed at a normal level, or overexpressed. In some embodiments, in order to determine the effect of an anti-tumor cell antigen therapeutic, a test biological sample comprising tumor cell antigen-expressing neoplastic cells is contacted with the anti-tumor cell antigen therapeutic agent for a sufficient time to allow the therapeutic agent to exert a cellular response, and then expression level of one or more biomarkers of interest in that test biological sample is compared to the expression level in the control biological sample in the absence of the anti-tumor cell antigen therapeutic agent. In some embodiments, the control biological sample of neoplastic cells is contacted with a neutral substance or negative control. For example, in some embodiments, a non-specific immunoglobulin, for example IgG1, which does not bind to tumor cell antigen serves as the negative control. Detection can occur over a time course to allow for

monitoring of changes in expression products over time. Detection can also occur with exposure to different concentrations of the anti-tumor cell antigen therapeutic agent to generate a “dose-response” curve for any given biomarker of interest.

Detection of cancer phenotype

[00288] Once expressed and, if necessary, purified, the cancer-associated proteins and nucleic acids are useful in a number of applications. In some embodiments, the expression levels of genes are determined for different cellular states in the cancer phenotype; that is, the expression levels of genes in normal tissue and in cancer tissue (and in some cases, for varying severities of lymphoma that relate to prognosis, as outlined below) are evaluated to provide expression profiles. An expression profile of a particular cell state or point of development is essentially a “fingerprint” of the state; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is unique to the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. Then, diagnosis may be done or confirmed: does tissue from a particular patient have the gene expression profile of normal or cancer tissue.

[00289] “Differential expression,” or equivalents used herein, refers to both qualitative as well as quantitative differences in the temporal and/or cellular expression patterns of genes, within and among the cells. Thus, a differentially expressed gene can qualitatively have its expression altered, including an activation or inactivation, in, for example, normal versus cancer tissue. That is, genes may be turned on or turned off in a particular state, relative to another state. As is apparent to the skilled artisan, any comparison of two or more states can be made. Such a qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard techniques in one such state or cell type, but is not detectable in both. Alternatively, the determination is quantitative in that expression is increased or decreased; that is, the expression of the gene is either up-regulated, resulting in an increased amount of transcript, or down-regulated, resulting in a decreased amount of transcript. The degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip® expression arrays, Lockhart, *Nature Biotechnology*, 14:1675-1680 (1996), hereby expressly incorporated by reference. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, Northern analysis and RNase protection. As outlined above, the change in expression (i.e. upregulation or

downregulation) is at least about 2-fold, 3-fold, 5-fold, 10-fold, 20-fold, 50-fold, or even 100 fold or more.

[00290] As will be appreciated by those in the art, this may be done by evaluation at either the gene transcript, or the protein level; that is, the amount of gene expression may be monitored using nucleic acid probes to the DNA or RNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) can be monitored, for example through the use of antibodies to the cancer-associated protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass spectroscopy assays, 2D gel electrophoresis assays, etc. Thus, the proteins corresponding to cancer-associated genes, i.e. those identified as being important in a particular cancer phenotype, i.e., lymphoma, can be evaluated in a diagnostic test specific for that cancer.

[00291] In some embodiments, gene expression monitoring is performed and a number of genes are monitored simultaneously. However, multiple protein expression monitoring can be done as well to prepare an expression profile. Alternatively, these assays may be done on an individual basis.

[00292] In some embodiments, the cancer-associated nucleic acid probes may be attached to biochips as outlined herein for the detection and quantification of cancer-associated sequences in a particular cell. The assays are done as is known in the art. As will be appreciated by those in the art, any number of different cancer-associated sequences may be used as probes, with single sequence assays being used in some cases, and a plurality of the sequences described herein being used in other embodiments. In addition, while solid-phase assays are described, any number of solution based assays may be done as well.

[00293] In some embodiments, both solid and solution based assays may be used to detect cancer-associated sequences that are up-regulated or down-regulated in cancers as compared to normal tissue. In instances where the cancer-associated sequence has been altered but shows the same expression profile or an altered expression profile, the protein will be detected as outlined herein.

[00294] In some embodiments nucleic acids encoding the cancer-associated protein are detected. Although DNA or RNA encoding the cancer-associated protein may be detected, of particular interest are methods wherein the mRNA encoding a cancer-associated protein is detected. The presence of mRNA in a sample is an indication that the cancer-associated gene has been transcribed to form the mRNA, and suggests that the protein is expressed. Probes to detect

the mRNA can be any nucleotide/deoxynucleotide probe that is complementary to and base pairs with the mRNA and includes but is not limited to oligonucleotides, cDNA or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be examined on a solid support such as nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is detected. In another method detection of the mRNA is performed in situ. In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following washing to remove the non-specifically bound probe, the label is detected. For example a digoxigenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding a cancer-associated protein is detected by binding the digoxigenin with an anti-digoxigenin secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

[00295] Any of the three classes of proteins as described herein (secreted, transmembrane or intracellular proteins) may be used in diagnostic assays. The cancer-associated proteins, antibodies, nucleic acids, modified proteins and cells containing cancer-associated sequences are used in diagnostic assays. This can be done on an individual gene or corresponding polypeptide level, or as sets of assays.

[00296] As described and defined herein, cancer-associated proteins find use as markers of cancers, including leukemia, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, skin cancer and lymphomas such as, but not limited to, Hodgkin's and non-Hodgkin's lymphoma. Detection of these proteins in putative cancer tissue or patients allows for a determination or diagnosis of the type of cancer. Numerous methods known to those of ordinary skill in the art find use in detecting cancers.

[00297] Antibodies may be used to detect cancer-associated proteins. One method separates proteins from a sample or patient by electrophoresis on a gel (typically a denaturing and reducing protein gel, but may be any other type of gel including isoelectric focusing gels and the like). Following separation of proteins, the cancer-associated protein is detected by immunoblotting with antibodies raised against the cancer-associated protein. Methods of immunoblotting are well known to those of ordinary skill in the art. The antibodies used in such methods may be labeled as described above.

[00298] In some methods, antibodies to the cancer-associated protein find use in in situ imaging techniques. In this method cells are contacted with from one to many antibodies to the cancer-associated protein(s). Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is detected. In some embodiments the antibody is detected by incubating with a secondary antibody that contains a detectable label. In another method the primary antibody to the cancer-associated protein(s) contains a detectable label. In another method, each one of multiple primary antibodies contains a distinct and detectable label. This method finds particular use in simultaneous screening for a plurality of cancer-associated proteins. As will be appreciated by one of ordinary skill in the art, numerous other histological imaging techniques are useful in the invention.

[00299] The label may be detected in a fluorometer that has the ability to detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) can be used in the method.

[00300] Antibodies may be used in diagnosing cancers from blood samples. As previously described, certain cancer-associated proteins are secreted/circulating molecules. Blood samples, therefore, are useful as samples to be probed or tested for the presence of secreted cancer-associated proteins. Antibodies can be used to detect the cancer-associated proteins by any of the previously described immunoassay techniques including ELISA, immunoblotting (Western blotting), immunoprecipitation, BIAcore technology and the like, as will be appreciated by one of ordinary skill in the art.

[00301] In situ hybridization of labeled cancer-associated nucleic acid probes to tissue arrays may be carried out. For example, arrays of tissue samples, including cancer-associated tissue and/or normal tissue, are made. In situ hybridization as is known in the art can then be done.

[00302] It is understood that when comparing the expression fingerprints between an individual and a standard, the skilled artisan can make a diagnosis as well as a prognosis. It is further understood that the genes that indicate diagnosis may differ from those that indicate prognosis.

[00303] As noted above, the cancer-associated proteins, antibodies, nucleic acids, modified proteins and cells containing cancer-associated sequences can be used in prognosis assays. As above, gene expression profiles can be generated that correlate to cancer severity, in terms of long term prognosis. Again, this may be done on either a protein or gene level. As above, the cancer-associated probes may be attached to biochips for the detection and quantification of cancer-associated sequences in a tissue or patient. The assays proceed as outlined for diagnosis.

Screening Assays

[00304] Any of the cancer-associated gene sequences as described herein may be used in drug screening assays. The cancer-associated proteins, antibodies, nucleic acids, modified proteins and cells containing cancer-associated gene sequences are used in drug screening assays or by evaluating the effect of drug candidates on a “gene expression profile” or expression profile of polypeptides. In one method, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, Zlokarnik, et al., Science 279, 84-8 (1998), Heid, et al., Genome Res., 6:986-994 (1996).

[00305] In some embodiments, the cancer associated proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified cancer associated proteins are used in screening assays. That is, the present invention provides novel methods for screening for compositions that modulate the cancer phenotype. As above, this can be done by screening for modulators of gene expression or for modulators of protein activity. Similarly, this may be done on an individual gene or protein level or by evaluating the effect of drug candidates on a “gene expression profile”. In some embodiments, the expression profiles are used, preferably sometimes in conjunction with high throughput screening techniques, to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, *supra*.

[00306] A variety of assays to evaluate the effects of agents on gene expression may be performed. In some embodiments, assays may be run on an individual gene or protein level. That is, candidate bioactive agents may be screened to modulate the gene's regulation. “Modulation” thus includes both an increase and a decrease in gene expression or activity. The amount of modulation will depend on the original change of the gene expression in normal versus tumor tissue, with changes of at least 10%, at least 50%, at least 100-300%, and at least 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in tumor compared to normal tissue, a decrease of about four fold may be desired; a 10-fold decrease in tumor compared to normal tissue gives a 10-fold increase in expression for a candidate agent is desired, *etc.* Alternatively, where the cancer-associated sequence has been altered but shows the same expression profile or an altered expression profile, the protein will be detected as outlined herein.

[00307] As will be appreciated by those in the art, this may be done by evaluation at either the gene or the protein level; that is, the amount of gene expression may be monitored using nucleic

acid probes and the quantification of gene expression levels, or, alternatively, the level of the gene product itself can be monitored, for example through the use of antibodies to the cancer-associated protein and standard immunoassays. Alternatively, binding and bioactivity assays with the protein may be done as outlined below.

[00308] In some some embodiments, a number of genes are monitored simultaneously, i.e. an expression profile is prepared, although multiple protein expression monitoring can be done as well.

[00309] In some embodiments, the cancer-associated nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of cancer-associated sequences in a particular cell. The assays are further described below.

[00310] In some embodiments a candidate bioactive agent is added to the cells prior to analysis. Moreover, screens are provided to identify a candidate bioactive agent that modulates a particular type of cancer, modulates cancer-associated proteins, binds to a cancer-associated protein, or interferes between the binding of a cancer-associated protein and an antibody.

[00311] The term "candidate bioactive agent" or "drug candidate" or grammatical equivalents as used herein describes any molecule, e.g., protein, oligopeptide, small organic or inorganic molecule, polysaccharide, polynucleotide, etc., to be tested for bioactive agents that are capable of directly or indirectly altering either the cancer phenotype, binding to and/or modulating the bioactivity of a cancer-associated protein, or the expression of a cancer-associated sequence, including both nucleic acid sequences and protein sequences. In some embodiments, the candidate agent suppresses a cancer-associated phenotype, for example to a normal tissue fingerprint. Similarly, the candidate agent preferably suppresses a severe cancer-associated phenotype. Generally a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

[00312] In some embodiments a candidate agent will neutralize the effect of a cancer-associated protein. By "neutralize" is meant that activity of a protein is either inhibited or counter acted against so as to have substantially no effect on a cell and hence reduce the severity of cancer, or prevent the incidence of cancer.

[00313] Candidate agents encompass numerous chemical classes, though typically they are organic or inorganic molecules, preferably small organic compounds having a molecular weight

of more than 100 and less than about 2,500 Daltons. In some embodiments small molecules are less than 2000, less than 1500, less than 1000, or less than 500 Da. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof.

[00314] Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides. In some embodiments libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, or amidification to produce structural analogs.

[00315] In some embodiments, the candidate bioactive agents are proteins. By "protein" herein is meant at least two covalently attached amino acids, which includes proteins, polypeptides, oligopeptides and peptides. The protein may be made up of naturally occurring amino acids and peptide bonds, or synthetic peptidomimetic structures. Thus "amino acid", or "peptide residue", as used herein means both naturally occurring and synthetic amino acids. For example, homophenylalanine, citrulline and norleucine are considered amino acids for the purposes of the invention. "Amino acid" also includes imino acid residues such as proline and hydroxyproline. The side chains may be in either the (R) or the (S) configuration. In some embodiments, the amino acids are in the (S) or L-configuration. If non-naturally occurring side chains are used, non-amino acid substituents may be used, for example to prevent or retard in vivo degradations.

[00316] In some embodiments, the candidate bioactive agents are naturally occurring proteins or fragments of naturally occurring proteins. Thus, for example, cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of prokaryotic and eukaryotic proteins may be made for screening in the methods

of the invention. In some embodiments the libraries are of bacterial, fungal, viral, and mammalian proteins. In some embodiments the library is a human proteinlibrary.

[00317] In some embodiments, the candidate bioactive agents are peptides of from about 5 to about 30 amino acids, from about 5 to about 20 amino acids, or from about 7 to about 15 amino acids. The peptides may be digests of naturally occurring proteins as is outlined above, random peptides, or “biased” random peptides. By “randomized” or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. Since generally these random peptides (or nucleic acids, discussed below) are chemically synthesized, they may incorporate any nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

[00318] In some embodiments, the library is fully randomized, with no sequence preferences or constants at any position. In some embodiments, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. For example, in some embodiments, the nucleotides or amino acid residues are randomized within a defined class, for example, of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc., or to purines, etc.

[00319] In some embodiments, the candidate bioactive agents are nucleic acids. As described generally for proteins, nucleic acid candidate bioactive agents may be naturally occurring nucleic acids, random nucleic acids, or “biased” random nucleic acids. In another embodiment, the candidate bioactive agents are organic chemical moieties, a wide variety of which are available in the literature.

[00320] In assays for testing alteration of the expression profile of one or more cancer-associated genes, after the candidate agent has been added and the cells incubated for some period of time, a nucleic acid sample containing the target sequences to be analyzed is prepared. The target sequence is prepared using known techniques (e.g., converted from RNA to labeled cDNA, as described above) and added to a suitable microarray. For example, an in vitro reverse transcription with labels covalently attached to the nucleosides is performed. In some

embodiments the nucleic acids are labeled with a label as defined herein, especially with biotin-FITC or PE, Cy3 and Cy5.

[00321] As will be appreciated by those in the art, these assays can be direct hybridization assays or can comprise “sandwich assays”, which include the use of multiple probes, as is generally outlined in U.S. Patent Nos. 5,681,702, 5,597,909, 5,545,730, 5,594,117, 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246 and 5,681,697, all of which are hereby incorporated by reference. In some embodiments, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

[00322] A variety of hybridization conditions may be used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions that allow formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration, pH, organic solvent concentration, etc. These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus, in some embodiments certain steps are performed at higher stringency conditions to reduce non-specific binding.

[00323] The reactions outlined herein may be accomplished in a variety of ways, as will be appreciated by those in the art. Components of the reaction may be added simultaneously, or sequentially, in any order, with suggested embodiments outlined below. In addition, the reaction may include a variety of other reagents in the assays. These include reagents like salts, buffers, neutral proteins, e.g. albumin, detergents, etc which may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used, depending on the sample preparation methods and purity of the target. In addition, either solid phase or solution based (i.e., kinetic PCR) assays may be used.

[00324] Once the assay is run, the data are analyzed to determine the expression levels, and changes in expression levels as between states, of individual genes, forming a gene expression profile.

[00325] In some embodiments, as for the diagnosis and prognosis applications, having identified the differentially expressed gene(s) or mutated gene(s) important in any one state, screens can be run to test for alteration of the expression of the cancer-associated genes individually. That is, screening for modulation of regulation of expression of a single gene can be done. Thus, for example, in the case of target genes whose presence or absence is unique between two states, screening is done for modulators of the target gene expression.

[00326] In addition, screens can be done for novel genes that are induced in response to a candidate agent. After identifying a candidate agent based upon its ability to suppress a cancer-associated expression pattern leading to a normal expression pattern, or modulate a single cancer-associated gene expression profile so as to mimic the expression of the gene from normal tissue, a screen as described above can be performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated cancer-associated tissue reveals genes that are not expressed in normal tissue or cancer-associated tissue, but are expressed in agent treated tissue. These agent specific sequences can be identified and used by any of the methods described herein for cancer-associated genes or proteins. In some embodiments these sequences and the proteins they encode find use in marking or identifying agent-treated cells. In addition, antibodies can be raised against the agent-induced proteins and used to target novel therapeutics to the treated cancer-associated tissue sample.

[00327] Thus, in some embodiments, a candidate agent is administered to a population of cancer-associated cells, that thus have an associated cancer-associated expression profile. By “administration” or “contacting” herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid encoding a proteinaceous candidate agent (i.e. a peptide) may be put into a viral construct such as a retroviral construct and added to the cell, such that expression of the peptide agent is accomplished; see PCT US97/01019, hereby expressly incorporated by reference.

[00328] Once the candidate agent has been administered to the cells, the cells can be washed if desired and are allowed to incubate under preferably physiological conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

[00329] Thus, for example, cancer-associated tissue may be screened for agents that reduce or suppress the cancer-associated phenotype. A change in at least one gene of the expression profile

indicates that the agent has an effect on cancer-associated activity. By defining such a signature for the cancer-associated phenotype, screens for new drugs that alter the phenotype can be devised. With this approach, the drug target need not be known and need not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.

[00330] In some embodiments, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of either the expression of the gene or the gene product itself can be done. The cancer-associated protein may be a fragment, or alternatively, be the full-length protein to the fragment encoded by the cancer-associated genes recited above. In some embodiments, the sequences are sequence variants as further described above.

[00331] In some embodiments the cancer-associated protein is a fragment approximately 14 to 24 amino acids in length. In some embodiments the fragment is a soluble fragment. In some embodiments, the fragment includes a non-transmembrane region. In some embodiments, the fragment has an N-terminal Cys to aid in solubility. In some embodiments, the C-terminus of the fragment is kept as a free acid and the N-terminus is a free amine to aid in coupling, e.g., to a cysteine.

[00332] In some embodiments the cancer-associated proteins are conjugated to an immunogenic agent as discussed herein. In some embodiments the cancer-associated protein is conjugated to BSA.

[00333] In some embodiments, screening is done to alter the biological function of the expression product of the cancer-associated gene. Again, having identified the importance of a gene in a particular state, screening for agents that bind and/or modulate the biological activity of the gene product can be run as is more fully outlined below.

[00334] In some embodiments, screens are designed to first find candidate agents that can bind to cancer-associated proteins, and then these agents may be used in assays that evaluate the ability of the candidate agent to modulate the cancer-associated protein activity and the cancer phenotype. Thus, as will be appreciated by those in the art, there are a number of different assays that may be run; binding assays and activity assays.

[00335] In some embodiments, binding assays are done. In general, purified or isolated gene product is used; that is, the gene products of one or more cancer-associated nucleic acids are

made. In general, this is done as is known in the art. For example, antibodies are generated to the protein gene products, and standard immunoassays are run to determine the amount of protein present. In some embodiments, cells comprising the cancer-associated proteins can be used in the assays.

[00336] Thus, in some embodiments, the methods comprise combining a cancer-associated protein and a candidate bioactive agent, and determining the binding of the candidate agent to the cancer-associated protein. Some embodiments utilize the human or mouse cancer-associated protein, although other mammalian proteins may also be used, for example for the development of animal models of human disease. In some embodiments, as outlined herein, variant or derivative cancer-associated proteins may be used.

[00337] In some embodiments of the methods herein, the cancer-associated protein or the candidate agent is non-diffusably bound to an insoluble support having isolated sample receiving areas (e.g. a microtiter plate, an array, etc.). The insoluble support may be made of any composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of screening. The surface of such supports may be solid or porous and of any convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, Teflon®, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples.

[00338] The particular manner of binding of the composition is not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition and is nondiffusable. Some methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation sequence when the protein is bound to the support), direct binding to “sticky” or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

[00339] In some embodiments, the cancer-associated protein is bound to the support, and a candidate bioactive agent is added to the assay. In some embodiments, the candidate agent is bound to the support and the cancer-associated protein is added. Novel binding agents include

specific antibodies, non-natural binding agents identified in screens of chemical libraries or peptide analogs. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

[00340] The determination of the binding of the candidate bioactive agent to the cancer-associated protein may be done in a number of ways. In some embodiments, the candidate bioactive agent is labeled, and binding determined directly. For example, this may be done by attaching all or a portion of the cancer-associated protein to a solid support, adding a labeled candidate agent (for example a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as is known in the art.

[00341] In some embodiments, only one of the components is labeled. For example, the proteins (or proteinaceous candidate agents) may be labeled at tyrosine positions using ^{125}I , or with fluorophores. Alternatively, more than one component may be labeled with different labels; using ^{125}I for the proteins, for example, and a fluorophore for the candidate agents.

[00342] In some embodiments, the binding of the candidate bioactive agent is determined through the use of competitive binding assays. In some embodiments, the competitor is a binding moiety known to bind to the target molecule (i.e. cancer-associated protein), such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding as between the bioactive agent and the binding moiety, with the binding moiety displacing the bioactive agent.

[00343] In some embodiments, the candidate bioactive agent is labeled. Either the candidate bioactive agent, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at any temperature which facilitates optimal activity, typically between 4 and 40°C. Incubation periods are selected for optimum activity, but may also be optimized to facilitate rapid high throughput screening. Typically between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

[00344] In some embodiments, the competitor is added first, followed by the candidate bioactive agent. Displacement of the competitor is an indication that the candidate bioactive

agent is binding to the cancer-associated protein and thus is capable of binding to, and potentially modulating, the activity of the cancer-associated protein. In some embodiments, either component can be labeled. Thus, for example, if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. In some embodiments, if the candidate bioactive agent is labeled, the presence of the label on the support indicates displacement.

[00345] In some embodiments, the candidate bioactive agent is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the bioactive agent is bound to the cancer-associated protein with a higher affinity. Thus, if the candidate bioactive agent is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the candidate agent is capable of binding to the cancer-associated protein.

[00346] In some embodiments, the methods comprise differential screening to identify bioactive agents that are capable of modulating the activity of the cancer-associated proteins. In this embodiment, the methods comprise combining a cancer-associated protein and a competitor in a first sample. A second sample comprises a candidate bioactive agent, a cancer-associated protein and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the cancer-associated protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the cancer-associated protein.

[00347] In some embodiments utilizes differential screening to identify drug candidates that bind to the native cancer-associated protein, but cannot bind to modified cancer-associated proteins. The structure of the cancer-associated protein may be modeled, and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect cancer-associated bioactivity are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

[00348] Positive controls and negative controls may be used in the assays. In some embodiments all control and test samples are performed in at least triplicate to obtain statistically significant results. Incubation of all samples is for a time sufficient for the binding of the agent to the protein. Following incubation, all samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is

employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

[00349] A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in any order that provides for the requisite binding.

[00350] Screening for agents that modulate the activity of cancer-associated proteins may also be done. In some embodiments, methods for screening for a bioactive agent capable of modulating the activity of cancer-associated proteins comprise adding a candidate bioactive agent to a sample of cancer-associated proteins, as above, and determining an alteration in the biological activity of cancer-associated proteins. "Modulating the activity of a cancer-associated protein" includes an increase in activity, a decrease in activity, or a change in the type or kind of activity present. Thus, in some embodiments, the candidate agent should both bind to cancer-associated proteins (although this may not be necessary), and alter its biological or biochemical activity as defined herein. The methods include both *in vitro* screening methods, as are generally outlined above, and *in vivo* screening of cells for alterations in the presence, distribution, activity or amount of cancer-associated proteins.

[00351] Thus, in some embodiments, the methods comprise combining a cancer-associated sample and a candidate bioactive agent, and evaluating the effect on cancer-associated activity. By "cancer-associated activity" or grammatical equivalents herein is meant one of the cancer-associated protein's biological activities, including, but not limited to, its role in tumorigenesis, including cell division, cell proliferation, tumor growth, cancer cell survival and transformation of cells. In some embodiments, cancer-associated activity includes activation of or by a protein encoded by a nucleic acid derived from a cancer-associated gene as identified above. An inhibitor of cancer-associated activity is the inhibitor of any one or more cancer-associated activities.

[00352] In some embodiments, the activity of the cancer-associated protein is increased; in some embodiments, the activity of the cancer-associated protein is decreased. Thus, bioactive agents are antagonists in some embodiments, and bioactive agents are agonists in some embodiments.

[00353] In some embodiments, the invention provides methods for screening for bioactive agents capable of modulating the activity of a cancer-associated protein. The methods comprise adding a candidate bioactive agent, as defined above, to a cell comprising cancer-associated proteins. Preferred cell types include almost any cell. The cells contain a recombinant nucleic acid that encodes a cancer-associated protein. In some embodiments, a library of candidate agents is tested on a plurality of cells.

[00354] In some embodiments, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, for example hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (i.e. cell-cell contacts). In some embodiments, the determinations are determined at different stages of the cell cycle process.

[00355] In this way, bioactive agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the cancer-associated protein.

Diagnosis and treatment of cancer

[00356] Methods of inhibiting cancer cell division are provided by the invention. In some embodiments, methods of inhibiting tumor growth are provided. In some embodiments, methods of treating cells or individuals with cancer are provided.

[00357] The methods may comprise the administration of a cancer inhibitor. In some embodiments, the cancer inhibitor is an antisense molecule, a pharmaceutical composition, a therapeutic agent or small molecule, or a monoclonal, polyclonal, chimeric or humanized antibody. In some embodiments, a therapeutic agent is coupled with an antibody. In some embodiments the therapeutic agent is coupled with a monoclonal antibody.

[00358] Methods for detection or diagnosis of cancer cells in an individual are also provided. In some embodiments, the diagnostic/detection agent is a small molecule that preferentially binds to a cancer-associated protein according to the invention. In some embodiments, the diagnostic/detection agent is an antibody

[00359] In some embodiments of the invention, animal models and transgenic animals are provided, which find use in generating animal models of cancers, particularly lymphoma, leukemia, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer or skin cancer. .

(a) Antisense molecules

[00360] The cancer inhibitor used may be an antisense molecule. Antisense molecules as used herein include antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for cancer molecules. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment generally of from about 14 to about 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen, *Cancer Res.* 48:2659, (1988) and van der Krol et al., *BioTechniques* 6:958, (1988).

[00361] Antisense molecules can be modified or unmodified RNA, DNA, or mixed polymer oligonucleotides. These molecules function by specifically binding to matching sequences resulting in inhibition of peptide synthesis (Wu-Pong, Nov 1994, *BioPharm*, 20-33) either by steric blocking or by activating an RNase H enzyme. Antisense molecules can also alter protein synthesis by interfering with RNA processing or transport from the nucleus into the cytoplasm (Mukhopadhyay & Roth, 1996, *Crit. Rev. in Oncogenesis* 7, 151-190). In addition, binding of single stranded DNA to RNA can result in nuclease-mediated degradation of the heteroduplex (Wu-Pong, *supra*). Backbone modified DNA chemistry which have thus far been shown to act as substrates for RNase H are phosphorothioates, phosphorodithioates, borontrifluoridates, and 2'-arabino and 2'-fluoro arabino-containing oligonucleotides.

[00362] Antisense molecules may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. In some embodiments, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

(b) RNA Interference

[00363] RNA interference refers to the process of sequence-specific post transcriptional gene silencing in animals mediated by short interfering RNAs (siRNA) (Fire et al., *Nature*, 391, 806 (1998)). The corresponding process in plants is referred to as post transcriptional gene silencing or RNA silencing and is also referred to as quelling in fungi. The presence of dsRNA in cells triggers the RNAi response through a mechanism that has yet to be fully characterized. This mechanism appears to be different from the interferon response that results from dsRNA mediated activation of protein kinase PKR and 2',5'-oligoadenylate synthetase resulting in non-specific cleavage of mRNA by ribonuclease L. (reviewed in Sharp, P.A., *RNA interference* – 2001, *Genes & Development* 15:485-490 (2001)).

[00364] Small interfering RNAs (siRNAs) are powerful sequence-specific reagents designed to suppress the expression of genes in cultured mammalian cells through a process known as RNA interference (RNAi). Elbashir, S.M. et al. *Nature* 411:494-498 (2001); Caplen, N.J. et al. *Proc. Natl. Acad. Sci. USA* 98:9742-9747 (2001); Harborth, J. et al. *J. Cell Sci.* 114:4557-4565 (2001). The term "short interfering RNA" or "siRNA" refers to a double stranded nucleic acid molecule capable of RNA interference "RNAi", (see Kreutzer et al., WO 00/44895; Zernicka-Goetz et al. WO 01/36646; Fire, WO 99/32619; Mello and Fire, WO 01/29058). As used herein, siRNA molecules are limited to RNA molecules but further encompasses chemically modified nucleotides and non-nucleotides. siRNA gene-targeting experiments have been carried out by transient siRNA transfer into cells (achieved by such classic methods as liposome-mediated transfection, electroporation, or microinjection).

[00365] Molecules of siRNA are 15- to 30-, 18- to 25-, or 21- to 23-nucleotide RNAs, with characteristic 2- to 3-nucleotide 3'-overhanging ends resembling the RNase III processing products of long double-stranded RNAs (dsRNAs) that normally initiate RNAi. When introduced into a cell, they assemble with yet-to-be-identified proteins of an endonuclease complex (RNA-induced silencing complex), which then guides target mRNA cleavage. As a consequence of degradation of the targeted mRNA, cells with a specific phenotype characteristic of suppression of the corresponding protein product are obtained. The small size of siRNAs, compared with traditional antisense molecules, prevents activation of the dsRNA-inducible interferon system present in mammalian cells. This avoids the nonspecific phenotypes normally produced by dsRNA larger than 30 base pairs in somatic cells.

[00366] Intracellular transcription of small RNA molecules is achieved by cloning the siRNA templates into RNA polymerase III (Pol III) transcription units, which normally encode the small

nuclear RNA (snRNA) U6 or the human RNase P RNA H1. Two approaches have been developed for expressing siRNAs: in the first, sense and antisense strands constituting the siRNA duplex are transcribed by individual promoters (Lee, N.S. et al. *Nat. Biotechnol.* 20, 500-505 (2002); Miyagishi, M. & Taira, K. *Nat. Biotechnol.* 20, 497-500 (2002).); in the second, siRNAs are expressed as fold-back stem-loop structures that give rise to siRNAs after intracellular processing (Paul, C.P. et al. *Nat. Biotechnol.* 20:505-508 (2002)). The endogenous expression of siRNAs from introduced DNA templates is thought to overcome some limitations of exogenous siRNA delivery, in particular the transient loss of phenotype. U6 and H1 RNA promoters are members of the type III class of Pol III promoters. (Paule, M.R. & White, R.J. *Nucleic Acids Res.* 28, 1283-1298 (2000)).

[00367] Co-expression of sense and antisense siRNAs mediate silencing of target genes, whereas expression of sense or antisense siRNA alone do not greatly affect target gene expression. Transfection of plasmid DNA, rather than synthetic siRNAs, may appear advantageous, considering the danger of RNase contamination and the costs of chemically synthesized siRNAs or siRNA transcription kits. Stable expression of siRNAs allows new gene therapy applications, such as treatment of persistent viral infections. Considering the high specificity of siRNAs, the approach also allows the targeting of disease-derived transcripts with point mutations, such as RAS or TP53 oncogene transcripts, without alteration of the remaining wild-type allele. Finally, by high-throughput sequence analysis of the various genomes, the DNA-based methodology may also be a cost-effective alternative for automated genome-wide loss-of-function phenotypic analysis, especially when combined with miniaturized array-based phenotypic screens. (Ziauddin, J. & Sabatini, D.M. *Nature* 411:107-110 (2001)).

[00368] The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNA) (Berstein et al., 2001, *Nature*, 409:363 (2001)). Short interfering RNAs derived from dicer activity are typically about 21-23 nucleotides in length and comprise about 19 base pair duplexes. Dicer has also been implicated in the excision of 21 and 22 nucleotide small temporal RNAs (stRNA) from precursor RNA of conserved structure that are implicated in translational control (Hutvagner et al., *Science*, 293, 834 (2001)). The RNAi response also features an endonuclease complex containing a siRNA, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single stranded RNA having sequence homologous to the siRNA. Cleavage of the target RNA

takes place in the middle of the region complementary to the guide sequence of the siRNA duplex (Elbashir et al., Genes Dev., 15, 188 (2001)).

[00369] The present invention provides expression systems comprising an isolated nucleic acid molecule comprising a sequence capable of specifically hybridizing to the cancer-associated sequences. In some embodiments, the nucleic acid molecule is capable of inhibiting the expression of the cancer-associated protein. A method of inhibiting expression of cancer-associated gene expression inside a cell by a vector-directed expression of a short RNA which short RNA can fold in itself and create a double strand RNA having cancer-associated mRNA sequence identity and able to trigger posttranscriptional gene silencing, or RNA interference (RNAi), of the cancer-associated gene inside the cell. In some embodiments a short double strand RNA having a cancer-associated mRNA sequence identity is delivered inside the cell to trigger posttranscriptional gene silencing, or RNAi, of the cancer-associated gene. In various embodiments, the nucleic acid molecule is at least a 7 mer, at least a 10 mer, or at least a 20 mer. In some embodiments, the sequence is unique.

[00370] In some embodiments the siRNA oligonucleotide has a sequence selected from the group consisting of SEQ ID NOS:1-144.

(c) Pharmaceutical Compositions

[00371] Pharmaceutical compositions encompassed by the present invention include as active agent, the polypeptides, polynucleotides, antisense oligonucleotides, or antibodies of the invention disclosed herein in a therapeutically effective amount. An "effective amount" is an amount sufficient to effect beneficial or desired results, including clinical results. An effective amount can be administered in one or more administrations. For purposes of this invention, an effective amount of an adenoviral vector is an amount that is sufficient to palliate, ameliorate, stabilize, reverse, slow or delay the progression of the disease state.

[00372] The compositions can be used to treat cancer as well as metastases of primary cancer. In addition, the pharmaceutical compositions can be used in conjunction with conventional methods of cancer treatment, e.g., to sensitize tumors to radiation or conventional chemotherapy. The terms "treatment", "treating", "treat" and the like are used herein to generally refer to obtaining a desired pharmacologic and/or physiologic effect. The effect may be prophylactic in terms of completely or partially preventing a disease or symptom thereof and/or may be therapeutic in terms of a partial or complete stabilization or cure for a disease and/or adverse effect attributable to the disease. "Treatment" as used herein covers any treatment of a disease in

a mammal, particularly a human, and includes: (a) preventing the disease or symptom from occurring in a subject which may be predisposed to the disease or symptom but has not yet been diagnosed as having it; (b) inhibiting the disease symptom, i.e., arresting its development; or (c) relieving the disease symptom, i.e., causing regression of the disease or symptom.

[00373] Where the pharmaceutical composition comprises an antibody that specifically binds to a gene product encoded by a differentially expressed polynucleotide, the antibody can be coupled to a drug for delivery to a treatment site or coupled to a detectable label to facilitate imaging of a site comprising cancer cells, such as prostate cancer cells. Methods for coupling antibodies to drugs and detectable labels are well known in the art, as are methods for imaging using detectable labels.

[00374] In some embodiments pharmaceutical compositions are provided comprising an antibody according to the present invention and a pharmaceutically suitable carrier, excipient or diluent. In some embodiments, the pharmaceutical composition further comprises a second therapeutic agent. In still another embodiment, the second therapeutic agent is a cancer chemotherapeutic agent.

[00375] A "patient" for the purposes of the present invention includes both humans and other animals, particularly mammals, and organisms. Thus the methods are applicable to both human therapy and veterinary applications. In some embodiments the patient is a mammal, and preferably the patient is human. One target patient population includes all patients currently undergoing treatment for cancer, particularly the specific cancer types mentioned herein. Subsets of these patient populations include those who have experienced a relapse of a previously treated cancer of this type in the previous six months and patients with disease progression in the past six months.

[00376] The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent to treat, ameliorate, or prevent a desired disease or condition, or to exhibit a detectable therapeutic or preventative effect. The effect can be detected by, for example, chemical markers or antigen levels. Therapeutic effects also include reduction in physical symptoms, such as decreased body temperature. The precise effective amount for a subject will depend upon the subject's size and health, the nature and extent of the condition, and the therapeutics or combination of therapeutics selected for administration. The effective amount for a given situation is determined by routine experimentation and is within the judgment of the clinician. For purposes of the present invention, an effective dose will generally be from about

0.01 mg/kg to about 5 mg/kg, from about 0.01 mg/kg to about 50 mg/kg, or about 0.05 mg/kg to about 10 mg/kg of the compositions of the present invention in the individual to which it is administered.

[00377] A pharmaceutical composition can also contain a pharmaceutically acceptable carrier. The term “pharmaceutically acceptable carrier” refers to a carrier for administration of a therapeutic agent, such as antibodies or a polypeptide, genes, and other therapeutic agents. The term refers to any pharmaceutical carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which can be administered without undue toxicity. Suitable carriers can be large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Pharmaceutically acceptable carriers in therapeutic compositions can include liquids such as water, saline, glycerol and ethanol. Auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, can also be present in such vehicles. In some embodiments, the therapeutic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection can also be prepared. Liposomes are included within the definition of a pharmaceutically acceptable carrier. Pharmaceutically acceptable salts can also be present in the pharmaceutical composition, e.g., mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in Remington: The Science and Practice of Pharmacy (1995) Alfonso Gennaro, Lippincott, Williams, & Wilkins.

[00378] The pharmaceutical compositions can be prepared in various forms, such as granules, tablets, pills, suppositories, capsules, suspensions, salves, lotions and the like. Pharmaceutical grade organic or inorganic carriers and/or diluents suitable for oral and topical use can be used to make up compositions containing the therapeutically-active compounds. Diluents known to the art include aqueous media, vegetable and animal oils and fats. Stabilizing agents, wetting and emulsifying agents, salts for varying the osmotic pressure or buffers for securing an adequate pH value, and skin penetration enhancers can be used as auxiliary agents.

[00379] The pharmaceutical compositions of the present invention comprise a cancer-associated protein in a form suitable for administration to a patient. In some embodiments, the pharmaceutical compositions are in a water soluble form, such as being present as

pharmaceutically acceptable salts, which is meant to include both acid and base addition salts. "Pharmaceutically acceptable acid addition salt" refers to those salts that retain the biological effectiveness of the free bases and that are not biologically or otherwise undesirable, formed with inorganic acids such as hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid and the like, and organic acids such as acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, maleic acid, malonic acid, succinic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluenesulfonic acid, salicylic acid and the like. "Pharmaceutically acceptable base addition salts" include those derived from inorganic bases such as sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts and the like. Particularly preferred are the ammonium, potassium, sodium, calcium, and magnesium salts. Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring substituted amines, cyclic amines and basic ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, and ethanolamine.

[00380] The pharmaceutical compositions may also include one or more of the following: carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose, lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol. Additives are well known in the art, and are used in a variety of formulations.

[00381] The compounds having the desired pharmacological activity may be administered in a physiologically acceptable carrier to a host, as previously described. The pharmaceutical compositions may be administered in a variety of routes including, but not limited to, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal or transcutaneous applications (for example, see WO98/20734), subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, intravaginal or rectal means. Depending upon the manner of introduction, the compounds may be formulated in a variety of ways. The concentration of therapeutically active compound in the formulation may vary from about 0.1-100% wgt/vol. Once formulated, the compositions contemplated by the invention can be (1) administered directly to the subject (e.g., as polynucleotide, polypeptides, small molecule agonists or antagonists, and the like); or (2) delivered ex vivo, to cells derived from the subject (e.g., as in ex vivo gene therapy). Direct delivery of the compositions will generally be accomplished by parenteral injection, e.g., subcutaneously, intraperitoneally, intravenously or

intramuscularly, intratumoral or to the interstitial space of a tissue. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal applications, needles, and gene guns (see the worldwideweb site at powderject.com) or hyposprays. Dosage treatment can be a single dose schedule or a multiple dose schedule.

[00382] Methods for the ex vivo delivery and reimplantation of transformed cells into a subject are known in the art and described in e.g., WO 93/14778. Examples of cells useful in ex vivo applications include, for example, stem cells, particularly hematopoietic, lymph cells, macrophages, dendritic cells, or tumor cells. Generally, delivery of nucleic acids for both ex vivo and in vitro applications can be accomplished by, for example, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei, all well known in the art.

[00383] Once differential expression of a gene corresponding to a cancer-associated polynucleotide described herein has been found to correlate with a proliferative disorder, such as neoplasia, dysplasia, and hyperplasia, the disorder can be amenable to treatment by administration of a therapeutic agent based on the provided polynucleotide, corresponding polypeptide or other corresponding molecule (e.g., antisense, ribozyme, etc.). In other embodiments, the disorder can be amenable to treatment by administration of a small molecule drug that, for example, serves as an inhibitor (antagonist) of the function of the encoded gene product of a gene having increased expression in cancerous cells relative to normal cells or as an agonist for gene products that are decreased in expression in cancerous cells (e.g., to promote the activity of gene products that act as tumor suppressors).

[00384] The dose and the means of administration of the inventive pharmaceutical compositions are determined based on the specific qualities of the therapeutic composition, the condition, age, and weight of the patient, the progression of the disease, and other relevant factors. For example, administration of polynucleotide therapeutic compositions agents includes local or systemic administration, including injection, oral administration, particle gun or catheterized administration, and topical administration. Preferably, the therapeutic polynucleotide composition contains an expression construct comprising a promoter operably linked to a polynucleotide of at least 12, 22, 25, 30, or 35 contiguous nt of the polynucleotide disclosed herein. Various methods can be used to administer the therapeutic composition directly to a specific site in the body. For example, a small metastatic lesion is located and the therapeutic composition injected several times in several different locations within the body of tumor.

Alternatively, arteries that serve a tumor are identified, and the therapeutic composition injected into such an artery, in order to deliver the composition directly into the tumor. A tumor that has a necrotic center is aspirated and the composition injected directly into the now empty center of the tumor. An antisense composition is directly administered to the surface of the tumor, for example, by topical application of the composition. X-ray imaging is used to assist in certain of the above delivery methods.

[00385] Targeted delivery of therapeutic compositions containing an antisense polynucleotide, subgenomic polynucleotides, or antibodies to specific tissues can also be used. Receptor-mediated DNA delivery techniques are described in, for example, Findeis et al., *Trends Biotechnol.* (1993) 11:202; Chiou et al., *Gene Therapeutics: Methods And Applications Of Direct Gene Transfer* (J.A. Wolff, ed.) (1994); Wu et al., *J. Biol. Chem.* (1988) 263:621; Wu et al., *J. Biol. Chem.* (1994) 269:542; Zenke et al., *Proc. Natl. Acad. Sci. (USA)* (1990) 87:3655; Wu et al., *J. Biol. Chem.* (1991) 266:338. Therapeutic compositions containing a polynucleotide are administered in a range of about 100 ng to about 200 mg of DNA for local administration in a gene therapy protocol. Concentration ranges of about 500 ng to about 50 mg, about 1 μ g to about 2 mg, about 5 μ g to about 500 μ g, and about 20 μ g to about 100 μ g of DNA can also be used during a gene therapy protocol. Factors such as method of action (e.g., for enhancing or inhibiting levels of the encoded gene product) and efficacy of transformation and expression are considerations that will affect the dosage required for ultimate efficacy of the antisense subgenomic polynucleotides. Where greater expression is desired over a larger area of tissue, larger amounts of antisense subgenomic polynucleotides or the same amounts re-administered in a successive protocol of administrations, or several administrations to different adjacent or close tissue portions of, for example, a tumor site, may be required to effect a positive therapeutic outcome. In all cases, routine experimentation in clinical trials will determine specific ranges for optimal therapeutic effect.

[00386] The therapeutic polynucleotides and polypeptides of the present invention can be delivered using gene delivery vehicles. The gene delivery vehicle can be of viral or non-viral origin (see generally, Jolly, *Cancer Gene Therapy* (1994) 1:51; Kimura, *Human Gene Therapy* (1994) 5:845; Connelly, *Human Gene Therapy* (1995) 1:185; and Kaplitt, *Nature Genetics* (1994) 6:148). Expression of such coding sequences can be induced using endogenous mammalian or heterologous promoters. Expression of the coding sequence can be either constitutive or regulated.

[00387] Viral-based vectors for delivery of a desired polynucleotide and expression in a desired cell are well known in the art. Exemplary viral-based vehicles include, but are not limited to, recombinant retroviruses (see, e.g., WO 90/07936; WO 94/03622; WO 93/25698; WO 93/25234; USPN 5,219,740; WO 93/11230; WO 93/10218; USPN 4,777,127; GB Patent No. 2,200,651; EP 0 345 242; and WO 91/02805), alphavirus-based vectors (e.g., Sindbis virus vectors, Semliki forest virus (ATCC VR-67; ATCC VR-1247), Ross River virus (ATCC VR-373; ATCC VR-1246) and Venezuelan equine encephalitis virus (ATCC VR-923; ATCC VR-1250; ATCC VR-1249; ATCC VR-532)), and adeno-associated virus (AAV) vectors (see, e.g., WO 94/12649, WO 93/03769; WO 93/19191; WO 94/28938; WO 95/11984 and WO 95/00655). Administration of DNA linked to killed adenovirus as described in Curiel, *Hum. Gene Ther.* (1992) 3:147 can also be employed.

[00388] Non-viral delivery vehicles and methods can also be employed, including, but not limited to, polycationic condensed DNA linked or unlinked to killed adenovirus alone (see, e.g., Curiel, *Hum. Gene Ther.* (1992) 3:147); ligand-linked DNA (see, e.g., Wu, *J. Biol. Chem.* (1989) 264:16985); eukaryotic cell delivery vehicles cells (see, e.g., US 5,814,482; WO 95/07994; WO 96/17072; WO 95/30763; and WO 97/42338) and nucleic charge neutralization or fusion with cell membranes. Naked DNA can also be employed. Exemplary naked DNA introduction methods are described in WO 90/11092 and US 5,580,859. Liposomes that can act as gene delivery vehicles are described in US 5,422,120; WO 95/13796; WO 94/23697; WO 91/14445; and EP 0524968. Additional approaches are described in Philip, *Mol. Cell Biol.* (1994) 14:2411, and in Woffendin, *Proc. Natl. Acad. Sci.* (1994) 91:1581.

[00389] Further non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Woffendin et al., *Proc. Natl. Acad. Sci. USA* (1994) 91(24):11581. Moreover, the coding sequence and the product of expression of such can be delivered through deposition of photopolymerized hydrogel materials or use of ionizing radiation (see, e.g., US 5,206,152 and WO 92/11033). Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun (see, e.g., US 5,149,655); use of ionizing radiation for activating transferred gene (see, e.g., USPN 5,206,152 and WO 92/11033).

[00390] In some embodiments, cancer-associated proteins and modulators are administered as therapeutic agents, and can be formulated as outlined above. Similarly, cancer-associated genes (including both the full-length sequence, partial sequences, or regulatory sequences of the cancer-associated coding regions) can be administered in gene therapy applications, as is known

in the art. These cancer-associated genes can include antisense applications, either as gene therapy (i.e. for incorporation into the genome) or as antisense compositions, as will be appreciated by those in the art.

[00391] Thus, in some embodiments, methods of modulating cancer-associated gene activity in cells or organisms are provided. In some embodiments, the methods comprise administering to a cell an anti-cancer-associated antibody that reduces or eliminates the biological activity of an endogenous cancer-associated protein. In some embodiments, the methods comprise administering to a cell or organism a recombinant nucleic acid encoding a cancer-associated protein. As will be appreciated by those in the art, this may be accomplished in any number of ways. In some embodiments, for example when the cancer-associated sequence is down-regulated in cancer, the activity of the cancer-associated expression product is increased by increasing the amount of cancer-associated expression in the cell, for example by overexpressing the endogenous cancer-associated gene or by administering a gene encoding the cancer-associated sequence, using known gene-therapy techniques. In some embodiments, the gene therapy techniques include the incorporation of the exogenous gene using enhanced homologous recombination (EHR), for example as described in PCT/US93/03868, hereby incorporated by reference in its entirety. In some embodiments, for example when the cancer-associated sequence is up-regulated in cancer, the activity of the endogenous cancer-associated gene is decreased, for example by the administration of a cancer-associated antisense nucleic acid.

(d) Vaccines

[00392] In some embodiments, cancer-associated genes are administered as DNA vaccines, either single genes or combinations of cancer-associated genes. Naked DNA vaccines are generally known in the art. Brower, Nature Biotechnology, 16:1304-1305 (1998).

[00393] In some embodiments, cancer-associated genes of the present invention are used as DNA vaccines. Methods for the use of genes as DNA vaccines are well known to one of ordinary skill in the art, and include placing a cancer-associated gene or portion of a cancer-associated gene under the control of a promoter for expression in a patient with cancer. The cancer-associated gene used for DNA vaccines can encode full-length cancer-associated proteins, but more preferably encodes portions of the cancer-associated proteins including peptides derived from the cancer-associated protein. In some embodiments a patient is immunized with a DNA vaccine comprising a plurality of nucleotide sequences derived from a

cancer-associated gene. Similarly, it is possible to immunize a patient with a plurality of cancer-associated genes or portions thereof. Without being bound by theory, expression of the polypeptide encoded by the DNA vaccine, cytotoxic T-cells, helper T-cells and antibodies are induced that recognize and destroy or eliminate cells expressing cancer-associated proteins.

[00394] In some embodiments, the DNA vaccines include a gene encoding an adjuvant molecule with the DNA vaccine. Such adjuvant molecules include cytokines that increase the immunogenic response to the cancer-associated polypeptide encoded by the DNA vaccine. Additional or alternative adjuvants are known to those of ordinary skill in the art and find use in the invention.

(e) Antibodies

[00395] The cancer-associated antibodies described above find use in a number of applications. For example, the cancer-associated antibodies may be coupled to standard affinity chromatography columns and used to purify cancer-associated proteins. The antibodies may also be used therapeutically as blocking polypeptides, as outlined above, since they will specifically bind to the cancer-associated protein.

[00396] The present invention further provides methods for detecting the presence of and/or measuring a level of a polypeptide in a biological sample, which cancer-associated polypeptide is encoded by a cancer-associated polynucleotide that is differentially expressed in a cancer cell, using an antibody specific for the encoded polypeptide. The methods generally comprise: a) contacting the sample with an antibody specific for a polypeptide encoded by a cancer-associated polynucleotide that is differentially expressed in a prostate cancer cell; and b) detecting binding between the antibody and molecules of the sample.

[00397] Detection of specific binding of the antibody specific for the encoded cancer-associated polypeptide, when compared to a suitable control is an indication that encoded polypeptide is present in the sample. Suitable controls include a sample known not to contain the encoded cancer-associated polypeptide or known not to contain elevated levels of the polypeptide; such as normal tissue, and a sample contacted with an antibody not specific for the encoded polypeptide, e.g., an anti-idiotypic antibody. A variety of methods to detect specific antibody-antigen interactions are known in the art and can be used in the method, including, but not limited to, standard immunohistological methods, immunoprecipitation, an enzyme immunoassay, and a radioimmunoassay. In general, the specific antibody will be detectably labeled, either directly or indirectly. Direct labels include radioisotopes; enzymes whose products are detectable (e.g.,

luciferase, β -galactosidase, and the like); fluorescent labels (e.g., fluorescein isothiocyanate, rhodamine, phycoerythrin, and the like); fluorescence emitting metals, e.g., ^{152}Eu , or others of the lanthanide series, attached to the antibody through metal chelating groups such as EDTA; chemiluminescent compounds, e.g., luminol, isoluminol, acridinium salts, and the like; bioluminescent compounds, e.g., luciferin, aequorin (green fluorescent protein), and the like. The antibody may be attached (coupled) to an insoluble support, such as a polystyrene plate or a bead. Indirect labels include second antibodies specific for antibodies specific for the encoded polypeptide ("first specific antibody"), wherein the second antibody is labeled as described above; and members of specific binding pairs, e.g., biotin-avidin, and the like. The biological sample may be brought into contact with and immobilized on a solid support or carrier, such as nitrocellulose, that is capable of immobilizing cells, cell particles, or soluble proteins. The support may then be washed with suitable buffers, followed by contacting with a detectably-labeled first specific antibody. Detection methods are known in the art and will be chosen as appropriate to the signal emitted by the detectable label. Detection is generally accomplished in comparison to suitable controls, and to appropriate standards.

[00398] In some embodiments, the methods are adapted for use *in vivo*, e.g., to locate or identify sites where cancer cells are present. In these embodiments, a detectably-labeled moiety, e.g., an antibody, which is specific for a cancer-associated polypeptide is administered to an individual (e.g., by injection), and labeled cells are located using standard imaging techniques, including, but not limited to, magnetic resonance imaging, computed tomography scanning, and the like. In this manner, cancer cells are differentially labeled.

(f) Other methods for the Detection and Diagnosis of Cancers

[00399] Without being bound by theory, the various cancer-associated sequences disclosed herein appear to be important in cancers. Accordingly, disorders based on mutant or variant cancer-associated genes may be determined. In some embodiments, the invention provides methods for identifying cells containing variant cancer-associated genes comprising determining all or part of the sequence of at least one endogenous cancer-associated genes in a cell. As will be appreciated by those in the art, this may be done using any number of sequencing techniques. In some embodiments, the invention provides methods of identifying the cancer-associated genotype of an individual comprising determining all or part of the sequence of at least one cancer-associated gene of the individual. This is generally done in at least one tissue of the individual, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced

cancer-associated gene to a known cancer-associated gene, i.e., a wild-type gene. As will be appreciated by those in the art, alterations in the sequence of some cancer-associated genes can be an indication of either the presence of the disease, or propensity to develop the disease, or prognosis evaluations.

[00400] The sequence of all or part of the cancer-associated gene can then be compared to the sequence of a known cancer-associated gene to determine if any differences exist. This can be done using any number of known homology programs, such as Bestfit, etc. In some embodiments, the presence of a difference in the sequence between the cancer-associated gene of the patient and the known cancer-associated gene is indicative of a disease state or a propensity for a disease state, as outlined herein.

[00401] In some embodiments, the cancer-associated genes are used as probes to determine the number of copies of the cancer-associated gene in the genome. For example, some cancers exhibit chromosomal deletions or insertions, resulting in an alteration in the copy number of a gene.

[00402] In some embodiments cancer-associated genes are used as probes to determine the chromosomal location of the cancer-associated genes. Information such as chromosomal location finds use in providing a diagnosis or prognosis in particular when chromosomal abnormalities such as translocations, and the like are identified in cancer-associated gene loci.

[00403] The present invention provides methods of using the polynucleotides described herein for detecting cancer cells, facilitating diagnosis of cancer and the severity of a cancer (e.g., tumor grade, tumor burden, and the like) in a subject, facilitating a determination of the prognosis of a subject, and assessing the responsiveness of the subject to therapy (e.g., by providing a measure of therapeutic effect through, for example, assessing tumor burden during or following a chemotherapeutic regimen). Detection can be based on detection of a polynucleotide that is differentially expressed in a cancer cell, and/or detection of a polypeptide encoded by a polynucleotide that is differentially expressed in a cancer cell. The detection methods of the invention can be conducted in vitro or in vivo, on isolated cells, or in whole tissues or a bodily fluid e.g., blood, plasma, serum, urine, and the like).

[00404] In some embodiments, methods are provided for detecting a cancer cell by detecting expression in the cell of a transcript that is differentially expressed in a cancer cell. Any of a variety of known methods can be used for detection, including, but not limited to, detection of a transcript by hybridization with a polynucleotide that hybridizes to a polynucleotide that is

differentially expressed in a prostate cancer cell; detection of a transcript by a polymerase chain reaction using specific oligonucleotide primers; in situ hybridization of a cell using as a probe a polynucleotide that hybridizes to a gene that is differentially expressed in a prostate cancer cell. The methods can be used to detect and/or measure mRNA levels of a gene that is differentially expressed in a cancer cell. In some embodiments, the methods comprise: a) contacting a sample with a polynucleotide that corresponds to a differentially expressed gene described herein under conditions that allow hybridization; and b) detecting hybridization, if any.

[00405] Detection of differential hybridization, when compared to a suitable control, is an indication of the presence in the sample of a polynucleotide that is differentially expressed in a cancer cell. Appropriate controls include, for example, a sample that is known not to contain a polynucleotide that is differentially expressed in a cancer cell, and use of a labeled polynucleotide of the same "sense" as the polynucleotide that is differentially expressed in the cancer cell. Conditions that allow hybridization are known in the art, and have been described in more detail above. Detection can also be accomplished by any known method, including, but not limited to, in situ hybridization, PCR (polymerase chain reaction), RT-PCR (reverse transcription-PCR), TMA, bDNA, and Nasbau and "Northern" or RNA blotting, or combinations of such techniques, using a suitably labeled polynucleotide. A variety of labels and labeling methods for polynucleotides are known in the art and can be used in the assay methods of the invention. Specificity of hybridization can be determined by comparison to appropriate controls.

[00406] Polynucleotides generally comprising at least 10 nt, at least 12nt or at least 15 contiguous nucleotides of a polynucleotide provided herein, are used for a variety of purposes, such as probes for detection of and/or measurement of, transcription levels of a polynucleotide that is differentially expressed in a prostate cancer cell. As will be readily appreciated by the ordinarily skilled artisan, the probe can be detectably labeled and contacted with, for example, an array comprising immobilized polynucleotides obtained from a test sample (e.g., mRNA). Alternatively, the probe can be immobilized on an array and the test sample detectably labeled. These and other variations of the methods of the invention are well within the skill in the art and are within the scope of the invention.

[00407] Nucleotide probes are used to detect expression of a gene corresponding to the provided polynucleotide. In Northern blots, mRNA is separated electrophoretically and contacted with a probe. A probe is detected as hybridizing to an mRNA species of a particular size. The amount of hybridization can be quantitated to determine relative amounts of expression, for example under a particular condition. Probes are used for in situ hybridization to

cells to detect expression. Probes can also be used in vivo for diagnostic detection of hybridizing sequences. Probes are typically labeled with a radioactive isotope. Other types of detectable labels can be used such as chromophores, fluorophores, and enzymes. Other examples of nucleotide hybridization assays are described in WO92/02526 and USPN 5,124,246.

[00408] PCR is another means for detecting small amounts of target nucleic acids (see, e.g., Mullis et al., *Meth. Enzymol.* (1987) 155:335; USPN 4,683,195; and USPN 4,683,202). Two primer oligonucleotides that hybridize with the target nucleic acids are used to prime the reaction. The primers can be composed of sequence within or 3' and 5' to the cancer-associated polynucleotides disclosed herein. Alternatively, if the primers are 3' and 5' to these polynucleotides, they need not hybridize to them or the complements. After amplification of the target with a thermostable polymerase, the amplified target nucleic acids can be detected by methods known in the art, e.g., Southern blot. mRNA or cDNA can also be detected by traditional blotting techniques (e.g., Southern blot, Northern blot, etc.) described in Sambrook et al., "Molecular Cloning: A Laboratory Manual" (New York, Cold Spring Harbor Laboratory, 1989) (e.g., without PCR amplification). In general, mRNA or cDNA generated from mRNA using a polymerase enzyme can be purified and separated using gel electrophoresis, and transferred to a solid support, such as nitrocellulose. The solid support is exposed to a labeled probe, washed to remove any unhybridized probe, and duplexes containing the labeled probe are detected.

[00409] Methods using PCR amplification can be performed on the DNA from a single cell, although it is convenient to use at least about 105 cells. The use of the polymerase chain reaction is described in Saiki et al. (1985) *Science* 239:487, and a review of current techniques may be found in Sambrook, et al. *Molecular Cloning: A Laboratory Manual*, CSH Press 1989, pp.14.2-14.33. A detectable label may be included in the amplification reaction. Suitable detectable labels include fluorochromes, (e.g. fluorescein isothiocyanate (FITC), rhodamine, Texas Red, phycoerythrin, allophycocyanin, 6-carboxyfluorescein (6-FAM), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein, 6-carboxy-X-rhodamine (ROX), 6-carboxy-2',4',7',4,7-hexachlorofluorescein (HEX), 5-carboxyfluorescein (5-FAM) or N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA)), radioactive labels, (e.g. 32P, 35S, 3H, etc.), and the like. The label may be a two stage system, where the polynucleotides is conjugated to biotin, haptens, etc. having a high affinity binding partner, e.g. avidin, specific antibodies, etc., where the binding partner is conjugated to a detectable label. The label may be conjugated to one or both of the

primers. Alternatively, the pool of nucleotides used in the amplification is labeled, so as to incorporate the label into the amplification product.

[00410] The reagents used in detection methods can be provided as part of a kit. Thus, the invention further provides kits for detecting the presence and/or a level of a polynucleotide that is differentially expressed in a cancer cell (e.g., by detection of an mRNA encoded by the differentially expressed gene of interest), and/or a polypeptide encoded thereby, in a biological sample. Procedures using these kits can be performed by clinical laboratories, experimental laboratories, medical practitioners, or private individuals. The kits of the invention for detecting a polypeptide encoded by a polynucleotide that is differentially expressed in a cancer cell may comprise a moiety that specifically binds the polypeptide, which may be an antibody that binds the polypeptide or fragment thereof. The kits of the invention used for detecting a polynucleotide that is differentially expressed in a prostate cancer cell may comprise a moiety that specifically hybridizes to such a polynucleotide. The kit may optionally provide additional components that are useful in the procedure, including, but not limited to, buffers, developing reagents, labels, reacting surfaces, means for detection, control samples, standards, instructions, and interpretive information.

[00411] The present invention further relates to methods of detecting/diagnosing a neoplastic or preneoplastic condition in a mammal (for example, a human). "Diagnosis" as used herein generally includes determination of a subject's susceptibility to a disease or disorder, determination as to whether a subject is presently affected by a disease or disorder, prognosis of a subject affected by a disease or disorder (e.g., identification of pre-metastatic or metastatic cancerous states, stages of cancer, or responsiveness of cancer to therapy), and therametrics (e.g., monitoring a subject's condition to provide information as to the effect or efficacy of therapy).

[00412] An "effective amount" is an amount sufficient to effect beneficial or desired results, including clinical results. An effective amount can be administered in one or more administrations.

[00413] A "cell sample" encompasses a variety of sample types obtained from an individual and can be used in a diagnostic or monitoring assay. The definition encompasses blood and other liquid samples of biological origin, solid tissue samples such as a biopsy specimen or tissue cultures or cells derived therefrom, and the progeny thereof. The definition also includes samples that have been manipulated in any way after their procurement, such as by treatment with reagents, solubilization, or enrichment for certain components, such as proteins or

polynucleotides. The term "cell sample" encompasses a clinical sample, and also includes cells in culture, cell supernatants, cell lysates, serum, plasma, biological fluid, and tissue samples.

[00414] As used herein, the terms "neoplastic cells", "neoplasia", "tumor", "tumor cells", "cancer" and "cancer cells", (used interchangeably) refer to cells which exhibit relatively autonomous growth, so that they exhibit an aberrant growth phenotype characterized by a significant loss of control of cell proliferation (i.e., de-regulated cell division). Neoplastic cells can be malignant or benign.

[00415] The terms "individual," "subject," "host," and "patient," are used interchangeably herein and refer to any mammalian subject for whom diagnosis, treatment, or therapy is desired, particularly humans. Other subjects may include cattle, dogs, cats, guinea pigs, rabbits, rats, mice, horses, and so on. Examples of conditions that can be detected/diagnosed in accordance with these methods include cancers. Polynucleotides corresponding to genes that exhibit the appropriate expression pattern can be used to detect cancer in a subject. For a review of markers of cancer, see, e.g., Hanahan et al. Cell 100:57-70 (2000).

[00416] In some embodiments detection/diagnostic methods comprise: (a) obtaining from a mammal (e.g., a human) a biological sample, (b) detecting the presence in the sample of a cancer-associated protein and (c) comparing the amount of product present with that in a control sample. In some embodiments, the presence in the sample of elevated levels of a cancer associated gene product indicates that the subject has a neoplastic or preneoplastic condition.

[00417] Biological samples suitable for use in this method include biological fluids such as serum, plasma, pleural effusions, urine and cerebro-spinal fluid, CSF, tissue samples (e.g., mammary tumor or prostate tissue slices) can also be used in the method of the invention, including samples derived from biopsies. Cell cultures or cell extracts derived, for example, from tissue biopsies can also be used.

[00418] In some embodiments the compound is a binding protein, e.g., an antibody, polyclonal or monoclonal, or antigen binding fragment thereof, which can be labeled with a detectable marker (e.g., fluorophore, chromophore or isotope, etc). Where appropriate, the compound can be attached to a solid support such as a bead, plate, filter, resin, etc. Determination of formation of the complex can be effected by contacting the complex with a further compound (e.g., an antibody) that specifically binds to the first compound (or complex). Like the first compound, the further compound can be attached to a solid support and/or can be labeled with a detectable marker.

[00419] The identification of elevated levels of cancer-associated protein in accordance with the present invention makes possible the identification of subjects (patients) that are likely to benefit from adjuvant therapy. For example, a biological sample from a post primary therapy subject (e.g., subject having undergone surgery) can be screened for the presence of circulating cancer-associated protein, the presence of elevated levels of the protein, determined by studies of normal populations, being indicative of residual tumor tissue. Similarly, tissue from the cut site of a surgically removed tumor can be examined (e.g., by immunofluorescence), the presence of elevated levels of product (relative to the surrounding tissue) being indicative of incomplete removal of the tumor. The ability to identify such subjects makes it possible to tailor therapy to the needs of the particular subject. Subjects undergoing non-surgical therapy, e.g., chemotherapy or radiation therapy, can also be monitored, the presence in samples from such subjects of elevated levels of cancer-associated protein being indicative of the need for continued treatment. Staging of the disease (for example, for purposes of optimizing treatment regimens) can also be effected, for example, by biopsy e.g. with antibody specific for a cancer-associated protein.

(g) Animal Models and Transgenics

[00420] The cancer-associated genes also find use in generating animal models of cancers, particularly lymphoma, leukemia, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer or skin cancer. As is appreciated by one of ordinary skill in the art, when the cancer-associated gene identified is repressed or diminished in cancer-associated tissue, gene therapy technology wherein antisense RNA directed to the cancer-associated gene will also diminish or repress expression of the gene. An animal generated as such serves as an animal model of cancer-associated that finds use in screening bioactive drug candidates. Similarly, gene knockout technology, for example as a result of homologous recombination with an appropriate gene targeting vector, will result in the absence of the cancer-associated protein. When desired, tissue-specific expression or knockout of the cancer-associated protein may be necessary.

[00421] It is also possible that the cancer-associated protein is overexpressed in cancer. As such, transgenic animals can be generated that overexpress the cancer-associated protein. Depending on the desired expression level, promoters of various strengths can be employed to express the transgene. Also, the number of copies of the integrated transgene can be determined and compared for a determination of the expression level of the transgene. Animals generated by such methods find use as animal models of cancer-associated and are additionally useful in screening for bioactive molecules to treat cancer.

Combination Therapy

[00422] In some embodiments the invention provides compositions comprising two or more cancer-associated gene antibodies to provide still improved efficacy against cancer.

Compositions comprising two or more cancer-associated gene antibodies may be administered to persons or mammals suffering from, or predisposed to suffer from, cancer. One or more cancer-associated gene antibodies may also be administered with another therapeutic agent, such as a cytotoxic agent, or cancer chemotherapeutic. Concurrent administration of two or more therapeutic agents does not require that the agents be administered at the same time or by the same route, as long as there is an overlap in the time period during which the agents are exerting their therapeutic effect. Simultaneous or sequential administration is contemplated, as is administration on different days or weeks.

[00423] In some embodiments the methods provide of the invention contemplate the administration of combinations, or "cocktails", of different antibodies. Such antibody cocktails may have certain advantages inasmuch as they contain antibodies which exploit different effector mechanisms or combine directly cytotoxic antibodies with antibodies that rely on immune effector functionality. Such antibodies in combination may exhibit synergistic therapeutic effects.

[00424] A cytotoxic agent refers to a substance that inhibits or prevents the function of cells and/or causes destruction of cells. The term is intended to include radioactive isotopes (e.g., I131, I125, Y90 and Re186), chemotherapeutic agents, and toxins such as enzymatically active toxins of bacterial, fungal, plant or animal origin or synthetic toxins, or fragments thereof. A non-cytotoxic agent refers to a substance that does not inhibit or prevent the function of cells and/or does not cause destruction of cells. A non-cytotoxic agent may include an agent that can be activated to be cytotoxic. A non-cytotoxic agent may include a bead, liposome, matrix or particle (see, e.g., U.S. Patent Publications 2003/0028071 and 2003/0032995 which are incorporated by reference herein). Such agents may be conjugated, coupled, linked or associated with an antibody according to the invention.

[00425] In some embodiments, conventional cancer medicaments are administered with the compositions of the present invention. Conventional cancer medicaments include:

- a) cancer chemotherapeutic agents.
- b) additional agents.
- c) prodrugs.

[00426] Cancer chemotherapeutic agents include, without limitation, alkylating agents, such as carboplatin and cisplatin; nitrogen mustard alkylating agents; nitrosourea alkylating agents, such as carmustine (BCNU); antimetabolites, such as methotrexate; folinic acid; purine analog antimetabolites, mercaptopurine; pyrimidine analog antimetabolites, such as fluorouracil (5-FU) and gemcitabine (Gemzar®); hormonal antineoplastics, such as goserelin, leuprolide, and tamoxifen; natural antineoplastics, such as aldesleukin, interleukin-2, docetaxel, etoposide (VP-16), interferon alfa, paclitaxel (Taxol®), and tretinoin (ATRA); antibiotic natural antineoplastics, such as bleomycin, dactinomycin, daunorubicin, doxorubicin, daunomycin and mitomycins including mitomycin C; and vinca alkaloid natural antineoplastics, such as vinblastine, vincristine, vindesine; hydroxyurea; aceglatone, adriamycin, ifosfamide, enocitabine, epitiostanol, aclarubicin, ancitabine, nimustine, procarbazine hydrochloride, carboquone, carboplatin, carmofur, chromomycin A3, antitumor polysaccharides, antitumor platelet factors, cyclophosphamide (Cytosin®), Schizophyllan, cytarabine (cytosine arabinoside), dacarbazine, thioinosine, thiotepa, tegafur, dolastatins, dolastatin analogs such as auristatin, CPT-11 (irinotecan), mitozantrone, vinorelbine, teniposide, aminopterin, carminomycin, esperamicins (See, e.g., U.S. Patent No. 4,675,187), neocarzinostatin, OK-432, bleomycin, furtulon, broxuridine, busulfan, honvân, peplomycin, bestatin (Ubenimex®), interferon- β , mepitiostane, mitobronitol, melphalan, laminin peptides, lentinan, Coriolus versicolor extract, tegafur/uracil, estramustine (estrogen/mechlorethamine).

[00427] Additional agents which may be used as therapy for cancer patients include EPO, G-CSF, ganciclovir; antibiotics, leuprolide; meperidine; zidovudine (AZT); interleukins 1 through 18, including mutants and analogues; interferons or cytokines, such as interferons α , β , and γ hormones, such as luteinizing hormone releasing hormone (LHRH) and analogues and, gonadotropin releasing hormone (GnRH); growth factors, such as transforming growth factor- β (TGF- β), fibroblast growth factor (FGF), nerve growth factor (NGF), growth hormone releasing factor (GHRF), epidermal growth factor (EGF), fibroblast growth factor homologous factor (FGFHF), hepatocyte growth factor (HGF), and insulin growth factor (IGF); tumor necrosis factor- α & β (TNF- α & β); invasion inhibiting factor-2 (IIF-2); bone morphogenetic proteins 1-7 (BMP 1-7); somatostatin; thymosin- α -1; γ -globulin; superoxide dismutase (SOD); complement factors; anti-angiogenesis factors; antigenic materials; and pro-drugs.

[00428] Prodrug refers to a precursor or derivative form of a pharmaceutically active substance that is less cytotoxic or non-cytotoxic to tumor cells compared to the parent drug and is capable of being enzymatically activated or converted into an active or the more active parent form. See,

e.g., Wilman, "Prodrugs in Cancer Chemotherapy" Biochemical Society Transactions, 14, pp. 375-382, 615th Meeting Belfast (1986) and Stella et al., "Prodrugs: A Chemical Approach to Targeted Drug Delivery," Directed Drug Delivery, Borchardt et al., (ed.), pp. 247-267, Humana Press (1985). Prodrugs include, but are not limited to, phosphate-containing prodrugs, thiophosphate-containing prodrugs, sulfate-containing prodrugs, peptide-containing prodrugs, D-amino acid-modified prodrugs, glycosylated prodrugs, b-lactam-containing prodrugs, optionally substituted phenoxyacetamide-containing prodrugs or optionally substituted phenylacetamide-containing prodrugs, 5-fluorocytosine and other 5-fluorouridine prodrugs which can be converted into the more active cytotoxic free drug. Examples of cytotoxic drugs that can be derivatized into a prodrug form for use herein include, but are not limited to, those chemotherapeutic agents described above.

Methods for delivering a cytotoxic agent or a diagnostic agent to a cell

[00429] The present invention also provides methods for delivering a cytotoxic agent or a diagnostic agent to one or more cells that express a cancer-associated gene. In some embodiments the methods comprise contacting an antibody, polypeptide or nucleotide of the present invention conjugated to a cytotoxic agent or diagnostic agent with the cell. Such conjugates are discussed above.

Affinity Purification

[00430] In some embodiments the invention provides methods and compositions for affinity purification. In some embodiments, antibodies of the invention are immobilized on a solid phase such as a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody is contacted with a sample containing the tumor cell antigen protein (or fragment thereof) to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the tumor cell antigen protein, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent, such as glycine buffer, pH 5.0, that will release the tumor cell antigen protein from the antibody.

EXAMPLES

[00431] The following examples are described so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the present invention, and are not intended to limit the scope of what the inventors regard as their invention nor are they intended to represent that the experiments below are all and only experiments performed. Efforts have been made to ensure accuracy with respect to numbers used (e.g. amounts, temperature,

etc.) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees Celsius, and pressure is at or near atmospheric.

Example 1: Insertion site analysis following tumor induction in mice

[00432] Tumors were induced in mice using either mouse mammary tumor virus (MMTV) or murine leukemia virus (MLV). MMTV causes mammary adenocarcinomas and MLV causes a variety of different hematopoietic malignancies (primarily T- or B-cell lymphomas).

[00433] Three routes of infection were used: (1) injection of neonates with purified virus preparations, (2) infection by milk-borne virus during nursing, and (3) genetic transmission of pathogenic proviruses via the germ-line (Akvr1 and/or Mtv2). The type of malignancy present in each affected mouse was determined by histological analysis of H&E-stained thin sections of formalin-fixed, paraffin-embedded biopsy samples. Host DNA sequences flanking all clonally-integrated proviruses in each tumor were recovered by nested anchored-PCR using two virus-specific primers and two primers specific for a 40 bp double stranded DNA anchor ligated to restriction enzyme digested tumor DNA. Amplified bands representing host/virus junction fragments were cloned and sequenced. Then the host sequences (called "tags") were used to BLAST analyze the mouse genomic sequence.

[00434] Extracted mouse genomic tag sequences were then mapped to the draft mouse genome assembly (NCBI m33 release) downloaded from www.ensembl.org. Tag sequences 45bp or longer were mapped to the genome using Timelogic's accelerated blast algorithm, terablast, with the following parameter setup: -t=10 -X=1e-10 -v=20 -b=20 -R. Short tag sequences (<45bp) were mapped to the genome by NCBI blastall algorithm, with the following parameter setup: -e 1000 -F F -W 9 -v 20 -b 20. The combined blast results were then filtered for the best matches for each tag sequence, which typically requires a minimum of 95% identity over at least 30% of the tag sequence length. Tags with uniq chromosome locations were passed on to the gene call process.

[00435] For each individual tag, three parameters were recorded: (1) the mouse chromosome assignment, (2) base pair coordinates at which the integration occurred, and (3) provirus orientation. Using this information, all available tags from all analyzed tumors were mapped to the mouse genome. To identify the protooncogene targets of provirus insertion mutation, the provirus integration pattern at each cluster of integrants was analyzed relative to the locations of all known genes in the transcriptome. The presence of provirus at the same locus in two or more

independent tumors is *prima facie* evidence that a protooncogene is present at or very near the proviral integration sites. This is because the genome is too large for random integrations to result in observable clustering. Any clustering that was detected provides unequivocal evidence for biological selection during tumorigenesis. In order to identify the human orthologs of the protooncogene targets of provirus insertion mutation, a comparative analysis of syntenic regions of the mouse and human genomes was performed.

[00436] Ensembl mouse gene models and UCSC refseq and knowngene sets were used to represent the mouse transcriptome. As noted above, based on the tag chromosome positions and the proviral insertion orientation relative to the adjacent genes, each tag was assigned to its nearest neighboring gene. Proviral insertions linked to a gene were grouped in 2 categories, type I insertions or type II insertions. If the insertion was within the gene locus, either intron or exon, it was designated as a type II insertion. If not, the insertion was designated as a type I insertion provided the insertion fulfilled these additional criteria: 1) it was outside the gene locus but within 100 kilobases from the gene's start or end positions, 2) for upstream insertions, the proviral orientation was the opposite to that of the gene, and 3) for downstream insertion, the proviral orientation was the same as the gene. Genes or transcripts discovered in this process were assigned with locus IDs from NCBI Locus Link annotations. The uniq mouse locus IDs with at least 2 viral inserts make up the current Oncogenome™.

[00437] To assign human orthologs for the mouse genes in the Oncogenome™, the MGI's mouse to human ortholog annotation and NCBI's homologue annotation was used. When there were conflicts or lack of ortholog annotation, comparative analysis of syntenic regions of the mouse and human genomes was performed, using the UCSC or Ensembl genome browser. The orthologous human genes were assigned with Locus Id's from NCBI Locus Link, and these human genes were further evaluated as potential targets for cancer therapeutics as described herein.

Example 2: Analysis of Quantitative RT-PCR: Comparative C_T Method.

[00438] The RT-PCR analysis was divided into 4 major steps: 1) RNA purification from primary normal and tumor tissues; 2) Generation of first strand cDNA from the purified tissue RNA for Real Time Quantitative PCR; 3) Setup RT-PCR for gene expression using ABI PRISM 7900HT Sequence Detection System tailored for 384-well reactions; 4) Analyze RT-PCR data by statistical methods to identify genes differentially expressed (up-regulated) in cancer.

[00439] These steps are set out in more detail below.

A) RNA purification from primary normal and tumor tissues

[00440] This was performed using Qiagen RNeasy mini Kit CAT#74106. Tissue chunks typically yielded approximately 30µg of RNA resulting in a final concentration of approximately 200 ng/µl if 150µl of elution buffer was used.

[00441] After RNA was extracted using Qiagen's protocol, Ribogreen quantitation reagents from Molecular Probes was used to determine yield and concentration of RNA according to manufacture protocol.

[00442] Integrity of extracted RNA was assessed on EtBr stained agarose gel to determine if the 28S and 18S band have equal intensity. In addition, sample bands should be clear and visible. If bands were not visible or smeared down through the gel, the sample was discarded.

[00443] Integrity of extracted RNA was also assessed using Agilent 2100 according to manufacture protocol. The Agilent Bioanalyzer / "Lab-On-A-Chip" is a micro-fluidics system that generates an electropherogram of an RNA sample. By observing the ratio of the 18S and 28S bands and the smoothness of the baseline a determination of the level of RNA degradation was made. Samples that have 28S: 18S ratio below 1 were discarded.

[00444] RNA samples were also examined by RT-PCR to determine level of genomic DNA contamination during extraction. In general, RNA samples were assayed directly using validated Taqman primers and probes of gene of interest in the presence and absence of Reverse Transcriptase. 12.5ng of RNA was used per reaction in quadruplicate in a 384 wells format in a volume of 5ul per well. (2ul of RNA + 3ul of RT+ or RT- master mix). The following thermocycle parameters was used (2-step PCR):

Thermocycling Parameters

Step	Reverse Transcription	Amp. Gold Activation	PCR	
	HOLD	HOLD	40 CYCLES	
			Denature	Anneal /Extend
Temperature	48°C	95°C	95°C	60°C
Time	30 min.	10 min.	15 sec.	1 min

[00445] RNA samples required the following criteria to consider as pass QC.

a) Ct difference must be 7 Ct or greater for a pass. Anything less is a "fail" and should be re-purified.

- b) Mean sample Ct must be within 2 STDEV (all samples) from Mean (all samples) to pass.
- c) Use conditional formatting to find the outliers of the sample group. *Do not include the outliers on the RNA panels.
- d) RT amplification or (Ct) must be > 34 cycles or it is a "fail".
- e) Human genomic DNA must be between 23 and 27.6 Ct.

[00446] RNA was assembled into panel only if samples passed all QC steps (Gel run, Agilent and RT-PCR for genomic DNA). RNA was arrayed for cDNA synthesis. In general, a minimum of 10 normals and 20 tumors were required for each tumor type (i.e., if a tissue type can have a squamous cell carcinoma and an adenocarcinoma, 20 samples of each tumor type must be used (the same 10 normals will be used for each tumor type)). In general, 11 µg of RNA was required per panel. A fudge factor of at least 2µg should be allowed; i.e., samples in database must have 13µg, or they will be dropped during cDNA array. Sample numbers were arranged in ascending orders, starting at well A1 and working down the column on 96 wells format. Four control samples will be placed at the end of the panel: hFB, hrRNA, hgDNA and Water (in that order). An additional NTC control (water) was be placed in well A2. All lot numbers of controls were recorded. RNA samples were normalised to 100ng/µl in Nuclease-free water. 11µg of RNA was used, the total volume being 110µl. NOTE: the concentration of RNA required can vary depending on the particular cDNA synthesis kit used. RNA samples that were below 100ng/µl, were loaded pure. After normalization was complete, the block was sealed using the heat sealer with easy peel foil @ 175°C for 2 seconds. The block was visually inspected to make sure foil was completely sealed. The manual sealer was then run over the foil. The block was stored in the -80°C freezers, ready for cDNA synthesis.

B) Generation of first strand cDNA from the purified tissue RNA for Real Time Quantitative PCR:

[00447] The following reaction mixture was setup in advance:

Reagents	1 RXN Volumes (μ l)	RXN
10X Taqman RT BUFFER	1	
25mM Magnesium chloride	2.2	
10mM deoxyNTPS mixture	2	
50uM Random Hexamer	0.5	
Rnase inhibitor	0.2	
50u/ul MultiScribe Rev. Transcriptase	0.25	
Water	0.85	

[00448] Arrayed RNA in a 96 well block (11 μ g) was distributed to daughter plates using Hydra to create 1 μ g of cDNA synthesis per 96 well plate. Each of these daughter plates was used to setup RT reaction using the following thermocycle parameters:

Step	Incubation	RT	RT Inactivation
	Hold	Hold	Hold
Time	10 min.	30 min.	5 min.
Temperature	25°C	48°C	95°C

[00449] Upon completion of thermocycling, plates were removed from the cycler and using the Hydra pipet, 60 μ l of 0.016M EDTA solution was pipetted into every well of cDNA the plates. Each cDNA plate (no more than 10 plates) was be pooled to a 2ml-96 well block for storage.

RT-PCR for gene expression using ABI PRISM 7900HT Sequence Detection System tailored for 384-well reactions:

Create Cocktails

[00450] Cocktails were produced as follows:

1. This protocol was designed to create cocktails for a panel with 96 samples; this is 470 rxns for the whole panel.
2. FRT (Forward and Reverse primers and Target probe) mix was removed from -20°C and place in 4°C fridge thaw.
3. The first 10 FRT's to be made were taken out and placed in a cold metal rack or in a rack on ice.
4. New 1.5ml cocktail tube caps were labelled with target number, side with the date of synthesis (found on FRT tube, if no date of synthesis label with today's date), and initials of scientist, one tube for each FRT being made.

RT tubes and cocktails tubes were organised in rack so that they were in order and easy to keep track of.

When pipeting a p200 was used at speed 6. Aspiration was carried out at the surface of the liquid, and dispensed near the top of the inside of the tube. Tips were changed after each aspirate/dispense step.

All cocktail tubes were opened and 94µl of Ambion water (poured fresh daily) was added, then tubes were closed.

The FRT was Pulse vortexed 15 times, then centrifuged for 10 sec. One by one 141µl of FRT was added to corresponding cocktail tubes.

When done with first 10, FRT was put back to -20°C immediately (if vol was less than 10µl then they were thrown away).

Cocktail was stored in 4°C until ready to run. (-20°C if it wait was longer than 1 day)

Master mix was added to cocktails when ready to run cocktails (refer to step 2.7)

7. Steps 1.3 to 1.6.5 were repeated for the next 10 cocktails, and so on until all cocktails had been made.

TaqMan Master Mix	1 rxn volume	470 RXNS	
TaqMan Universal Master Mix	2.5µl	1175µl	
Lot#			
Forward Primer working stock	0.1 µl	47 µl	} 141µl
Reverse Primer working stock	0.1 µl	47 µl	
Probe working stock	0.1 µl	47 µl	
Water	0.2 µl	94 µl	
Final Volume	3.0 µl	1410 µl	

[00451] 2µl of cDNA from the arrayed 96-well plates was added to the 3µl of Taqman Master Mix to makeup a 5µl QPCR reaction.

[00452] The primers and probes used in the QPCR for each gene are given in Table 2.

Table 2.

Gene	Primer Probe	ForwardPrimer	ReversePrimer	ProbeSequence
------	--------------	---------------	---------------	---------------

	ID			
WNT3	885	CGCACGCGTGGTAAATGAC (SEQ ID NO: 1)	GTAAGATGAGAGAGAGGGA GAGCCT (SEQ ID NO: 73)	CCCAACTCGCCTGTGGACGGG (SEQ ID NO: 145)
NOTCH1	91	AGTGTGCGTCCAACCCATG T (SEQ ID NO: 2)	CTGTGTAGGGCAGCAGGCA (SEQ ID NO: 74)	CACGTGTATTGACGACGTTGC CGG (SEQ ID NO: 146)
WNT1	44	CTGTGCGAGAAACGGCGTTT AT (SEQ ID NO: 3)	GCGCCACCGAATGGGT (SEQ ID NO: 75)	TCGCTATCACCTCCGCCGGG (SEQ ID NO: 147)
WNT3A	489	CTTCAAGGTGCCACGGA (SEQ ID NO: 4)	GGTTGGGCTCGCAGAAGTT (SEQ ID NO: 76)	CCTGGTCTACTACGAGGCCTC GCCC (SEQ ID NO: 148)
IL2RA	107	CTCTGTGACGATGACCCGC (SEQ ID NO: 5)	GGTTCCTTCCTTGTAGGCCA T (SEQ ID NO: 77)	AGAGATCCACACGCCACATT CAAAGC (SEQ ID NO: 149)
FLT3	160	TTTCAGTTTTTACCCCACTT TCCA (SEQ ID NO: 6)	CGAGTCCGGGTGTATCTGAA C (SEQ ID NO: 78)	ATCCAAATTCCAGCATGCCTG GTTCAAG (SEQ ID NO: 150)
PRLR	133	TGGGCAGCAAACAGAGTTT AAGA (SEQ ID NO: 7)	GGTTTGCAGCGAACCTGG (SEQ ID NO: 79)	TCTCAGCCTACATCCAGGACA GAAATACCTTG (SEQ ID NO: 151)
TNFRSF6	59	CAACAACCATGCTGGGCAT (SEQ ID NO: 8)	GGGCATTAACACTTTTGAC G (SEQ ID NO: 80)	TGGACCCTCCTACCTCTGGTTC TTACGTCT (SEQ ID NO: 152)
PTPRK	152	GGACTGGGAATCCTTGGGT TA (SEQ ID NO: 9)	TTGTGACCACGGAAGTAAT GGTA (SEQ ID NO: 81)	CGCGTTGCCACTCTTTAATGT CACTATCTG (SEQ ID NO: 153)
TNFSF11	881	TGGATGGCTCATGGTTAGA TCTG (SEQ ID NO: 10)	TGGCATTAAATAGTGAGATG AGCAAA (SEQ ID NO: 82)	TGAGCTTCAAGCTTGCTCCTCT TGGC (SEQ ID NO: 154)
FGFR3	800	TCCTGCTCTGGGAGATCTT CAC (SEQ ID NO: 11)	GGCCCTCCTTCAGCAGCTT (SEQ ID NO: 83)	CTCCTCCACAGGGATGCCGGG (SEQ ID NO: 155)
SELL	221	GAACTGGCAAAGGGCTAG AAGA (SEQ ID NO: 12)	TTCCGCCTTGTTTGTATG G (SEQ ID NO: 84)	CTGCCGAGACAATTACACAGA TTTAGTT (SEQ ID NO: 156)
DPT	66	TTGAGTCAGTGCTGGATC G (SEQ ID NO: 13)	CAGCAGGAATATGGGCACC T (SEQ ID NO: 85)	AGTGGCAGTTTACTGTTGTC GCTACAGCAA (SEQ ID NO: 157)
CD28	169	GAGGGATTCCCTCAAAGCA AT (SEQ ID NO: 14)	GTGAACATTAACACTGAAG TCTCATGAC (SEQ ID NO: 86)	CAGGTAAACCAAGTTGCTTTC CTCACTCCC (SEQ ID NO: 158)
PDCD1LG1	492	GCACTTTTCAAATGCCTTT GG (SEQ ID NO: 15)	GCCAGGTGTTGGCTAGTCTT GT (SEQ ID NO: 87)	TGACCCAGCACTTTAATCTGA AACCTGCA (SEQ ID NO: 159)
LFNG	178	TGACGCTGAGCTACGGTAT GTT (SEQ ID NO: 16)	CGAGAAGGGCCCCTTCAC (SEQ ID NO: 88)	AAAACAAGCGGAACGCCGTC CA (SEQ ID NO: 160)
ANXA2	289	GCCTTCCTGTCTAGTCTCTC CTGTA (SEQ ID NO: 17)	GAGGCAAAGTGTTTCACATC ATAGA (SEQ ID NO: 89)	CCAACTCCTTGGAATGTTTATT TCTTTGGC (SEQ ID NO: 161)
CDH23	521	GAAAGTCAATGTCAATGAA CAAGCA (SEQ ID NO: 18)	GGAACATCAGTCACTTCTGT TTATAAGC (SEQ ID NO: 90)	TTCAGTGGCTTCCCAAATGTGT GCC (SEQ ID NO: 162)
IFNAR2	886	ATCATTTTGTGACCTCACA GATGAG (SEQ ID NO: 19)	TGTTCCCGCTGAATCCTTCT AG (SEQ ID NO: 91)	AGCACACACGAGGCCTATGTC ACCG (SEQ ID NO: 163)
RAMP1		(SEQ ID NO: 20)	(SEQ ID NO: 92)	(SEQ ID NO: 164)
MET	76	TGCTAAAATGCTGGCACCC T (SEQ ID NO: 21)	GATATCCGGGACACCAGTT CA (SEQ ID NO: 93)	AGCCGAAATGCGCCCATCCTT T (SEQ ID NO: 165)
FGFR1	496	GGAGTGGGAGCCAATGAA CA (SEQ ID NO: 22)	CAGACCAAACCGACAGGAG AA (SEQ ID NO: 94)	AGCTCAGGAAGCTCTCACTTG CATGCC (SEQ ID NO: 166)
PLAU	497	CCCGACGGTGGGCATT (SEQ ID NO: 23)	CCTCAAGAGACCTCAGCTGT TACAC (SEQ ID NO: 95)	TGGGAAATTATTCACTTCTCA ACCATGGGC (SEQ ID NO: 167)
SLC4A8	1152	GGGCTTGCTGCTGGACAT (SEQ ID NO: 24)	TGAGTGCATCTCGGTAGTCG C (SEQ ID NO: 96)	AAGCGGAAGGCCCTGGTAC TG (SEQ ID NO: 168)

TPCN1	524	GGCCCATGCCACAGGTT (SEQ ID NO: 25)	ACCCTGGAGACCGGGAGTT (SEQ ID NO: 97)	TGTGTTTTGCTTTAGGGACAG AACCACCTTAGGAA (SEQ ID NO: 169)
LTBR	253	GCAGGTGGGCACTGGCT (SEQ ID NO: 26)	CTCAGGTTTGCTCAGGCAGT AG (SEQ ID NO: 98)	TCCTGTGGAGGGCACC GTACC C (SEQ ID NO: 170)
IL4R	254	CCTTCGAGCAGCACCTCCT (SEQ ID NO: 27)	GCAGACGGCCAGGATGAC (SEQ ID NO: 99)	AATGCAGGAAACGCTGACGCC C (SEQ ID NO: 171)
ITGA9	257	TGGAGCCCCGTGGCTAC (SEQ ID NO: 28)	GAGGATTCCCACCAACAAA CTG (SEQ ID NO: 100)	ATGGCGATGATCCACCCACG A (SEQ ID NO: 172)
FGFR2	258	CCACCCGGGATGGAGTACT (SEQ ID NO: 29)	GACACCAAGTCCTTGAAGG TCATC (SEQ ID NO: 101)	CCTATGACATTAACCGTGITC CTGAGGAGCA (SEQ ID NO: 173)
FCMD	423	GCTGTGGCCTTTCGGAAGA (SEQ ID NO: 30)	GAATGGTACTCCCAATTTGT TTAATGT (SEQ ID NO: 102)	TTCGCTGCTAGTTGCAGTAATT CCTTTGCA (SEQ ID NO: 174)
COL15A1	898	CGGGTGCTTTCTTCATTG AC (SEQ ID NO: 31)	TGCTTCCACATCAGTGATGG A (SEQ ID NO: 103)	TTTTCCCGCTAGTTCTCTTA (SEQ ID NO: 175)
SLAMF6	266	TCAATGAAACATCTCTTGC CTTCA (SEQ ID NO: 32)	CTTCCCTGTTTCGGATTAG TCA (SEQ ID NO: 104)	CCCCATGAAACAAAAGTCCA GAAATCC (SEQ ID NO: 176)
CD74	754	CCCCACACAGCTACAGCTT TCT (SEQ ID NO: 33)	ATGAGGTACAGGGTGGGAG ATG (SEQ ID NO: 105)	TCAGCCCCCAGCCCCCTCC (SEQ ID NO: 177)
TSPAN-2	269	GCCGTCATTGCTTTTGGAC TA (SEQ ID NO: 34)	GGGACTTGTCCTCTGATGAT AACTC (SEQ ID NO: 106)	ATGGCACCTCCGAACCGAAAC CA (SEQ ID NO: 178)
THSD2	461	GTGGCTTCAAAAGAGGGA CTGA (SEQ ID NO: 35)	TTTGCACTGTACACTTTCTT GTCTCA (SEQ ID NO: 107)	CGGGTCCGAGAAATAATACAG CATCCTTCA (SEQ ID NO: 179)
CD2	511	GGACATCTATCTCATCATT GGCATAT (SEQ ID NO: 36)	CTTTTGGTGATATAGAAAAC GAGCA (SEQ ID NO: 108)	TGCCACAAAGACCATCAAGAG GCTGC (SEQ ID NO: 180)
LNPEP	273	CGAACACAGAAGCTGTCTT TTATCA (SEQ ID NO: 37)	TGACAAAATCCCATGCCAG TAAG (SEQ ID NO: 109)	TGTCCAGGAAAATGTGACCC ACTGTTC (SEQ ID NO: 181)
IGSF8	536	GAGAAGGTGGCATCCAGA ACAT (SEQ ID NO: 38)	GGTAGGTGCCCGCATCAC (SEQ ID NO: 110)	ACCGGCTACGGCTAGAGGCTG CC (SEQ ID NO: 182)
ATP1A2	537	ACATGTATGCTTCCAGGGA CAAG (SEQ ID NO: 39)	ACCCAGAAAAAGCAAATCT CTTAGTATT (SEQ ID NO: 111)	TGTCCAAAGGTTCTGACCATG TTCATGACC (SEQ ID NO: 183)
ITGA4	552	GAGACTTCAGCACTCAAGT TTGAAA (SEQ ID NO: 40)	GCAACATTCTCATCCTTGTT TAGTTC (SEQ ID NO: 112)	AGCAACAGGTTTTCCAGAGCC AAATCCA (SEQ ID NO: 184)
B2M	883	CATTGCGGGCCGAGATGTCT (SEQ ID NO: 41)	AGGCCAGAAAGAGAGAGTA GCG (SEQ ID NO: 113)	CGCTCCGTGGCCTTAGCTGTG C (SEQ ID NO: 185)
PAG	560	TCCCAAAGGAGAACGACT ACGA (SEQ ID NO: 42)	TGGGTTGCTAGAGCCTGGTA ATA (SEQ ID NO: 114)	TCTCTGCCTTGCTGCAAGTCA CTTATGCT (SEQ ID NO: 186)
NOTCH2	888	GGAGGCTATGGCTGTGTAT GTGT (SEQ ID NO: 43)	AAGGCACAATCATCAATGTT CTCA (SEQ ID NO: 115)	TGCAGTCATCTCCACTCCAGC CGTT (SEQ ID NO: 187)
UTS2	804	CCGGAGGAGCTAGAAAGA GCTT (SEQ ID NO: 44)	GCTTTCCTGAGAATATCCCC TCTT (SEQ ID NO: 116)	CCCTTCTACAGATACTGCCAG AGATGCTGG (SEQ ID NO: 188)
TNFRSF14	667	CAGCAGCTCCCACTGGGTA (SEQ ID NO: 45)	TGGAGCAAACAATGACGAT GA (SEQ ID NO: 117)	CGAGGCTCCCTGAGAGAAACC ACCA (SEQ ID NO: 189)
TGFBR3	1112	TGAAGCTGCGCATGGTCTT A (SEQ ID NO: 46)	TCCTTGTGTGTGCTGGCATC (SEQ ID NO: 118)	CATCAGGCTGTGCAGGAAGAA GCTATCTG (SEQ ID NO: 190)
FGF8	32	GCAGAGTTCGAGTCCGAGG A (SEQ ID NO: 47)	CGATCAGCTTCCCCTTCTTG (SEQ ID NO: 119)	CGAGACGGGCCTCTACATCTG CATG (SEQ ID NO: 191)
FGF3	40	CCTTTTGTGGCGAACCGT (SEQ ID NO: 48)	GGAACGCGAGTCCCTTTAAT T (SEQ ID NO: 120)	CGCTTTGCGCAGCAACGGAA (SEQ ID NO: 192)

XYLT1	922	TGCCTACATGGAGCAGAGC TT (SEQ ID NO: 49)	CTGTTCCACCTGGGCGG (SEQ ID NO: 121)	CAGAGCCTAAACCCCGTCTC AGCC (SEQ ID NO: 193)
FGF4	900	CCAACAACCTACAACGCCTA CGA (SEQ ID NO: 50)	TTCTTGCTCAGGGCGATGA (SEQ ID NO: 122)	CCTACAAGTACCCCGGCAT (SEQ ID NO: 194)
LOC390243	791	GGACAGTCATGCCACCTG (SEQ ID NO: 51)	TCTTGTTGGTGTGGCATT A (SEQ ID NO: 123)	TGCAGATGTTGAGCAGCTCGT CCC (SEQ ID NO: 195)
THY1	638	GAAATGGCTTTTCCCAACC C (SEQ ID NO: 52)	GTGGTGGCACTATACACAGT CACA (SEQ ID NO: 124)	CCCACCCGGTGGTTCTTCCTGT TC (SEQ ID NO: 196)
SLC9A9	997	AAGGGTTATGTCAGAAAA GGATGAGT (SEQ ID NO: 53)	GCAAAAAATTGAAGACAAG CAGC (SEQ ID NO: 125)	TCAGTTTCAACATCAGGGAGC GGTGG (SEQ ID NO: 197)
PRG1	647	CGCTGCAATCCAGACAGTA ATTG (SEQ ID NO: 54)	GATTCACCTGGAAGTAGTTC GAACA (SEQ ID NO: 126)	TGCAAACTGCCTTGAAGAAAA AGGACCA (SEQ ID NO: 198)
CPM	957	GGTCAATTAAGAAAGCCCC TCTC (SEQ ID NO: 55)	TAATGGAGGACTGTGGCAT CC (SEQ ID NO: 127)	CTCCAGAGCCGTTTCCCTAGC TTTTGG (SEQ ID NO: 199)
FBN2	805	CAGAGGTCGATGAGGAAA ATGC (SEQ ID NO: 56)	TGCCTGCTGTCTTCTTAGA ATAGC (SEQ ID NO: 128)	TCCCCAGAAGCATGCTACGAG TGCA (SEQ ID NO: 200)
LTB	811	CAGAAGCTGCCAGAGGAG GA (SEQ ID NO: 57)	GAGGTGGGCAGCTGGGA (SEQ ID NO: 129)	CCAGAAACAGATCTCAGCCCC GGG (SEQ ID NO: 201)
KCNH5	816	CGGGCACTGACGTACTGTG A (SEQ ID NO: 58)	GCAAAAAGCTGTATAAAAGT CCAGGA (SEQ ID NO: 130)	ACATCATCAAGCGGGAAGCCT TGCT (SEQ ID NO: 202)
SLC12A7	891	CCGCCTACGCCAGACAAG (SEQ ID NO: 59)	GCTCCTGTACTTCTCAGCGA TCA (SEQ ID NO: 131)	TCTCCCTGGTCCAGGTCATCT GCAC (SEQ ID NO: 203)
FGF10	825	TGCATCATTTAACTGGCAG CATA (SEQ ID NO: 60)	CTCTCCTTGGAGCTCCTTTT CC (SEQ ID NO: 132)	TTCAATGCCACATACATTTGC CTCCCA (SEQ ID NO: 204)
IL7R	830	GCTTAGCCTTGGGACTACA AACAG (SEQ ID NO: 61)	GGTCAATGTCAGGATTCCA GATT (SEQ ID NO: 133)	ACGCTGCCCCCTCCATTTCTC TC (SEQ ID NO: 205)
MATN2	831	CCACACAGAGTTCACTCTG AGAACT (SEQ ID NO: 62)	GAGCCCTTTCCCATGTATTT CAT (SEQ ID NO: 134)	CCAAAGACATGAAAAAAGCC GTGGCC (SEQ ID NO: 206)
COL19A1	832	GCTGCCAGCAGCAATGTTG (SEQ ID NO: 63)	CCCAGGCAACCCATCCTT (SEQ ID NO: 135)	TGGTGGCCAGGTCTCCCAT AGC (SEQ ID NO: 207)
SLC1A2	847	CAAGTCTGAGCTGGATACC ATTGA (SEQ ID NO: 64)	TCATCATAAATGGATTGAGT CTTGGT (SEQ ID NO: 136)	TTCAATATCTTCATGCACTCGA TGCTGGG (SEQ ID NO: 208)
SLC28A2	892	GAGATGGAGAAAGCAAGT GGAAGA (SEQ ID NO: 65)	CCAGCCCCAGGTTCACTGT (SEQ ID NO: 137)	CAGTCCATTGCTCTGTCCACA GTGGAGA (SEQ ID NO: 209)
SLPI	849	AGGTCTTTCCACCCTGAG ACT (SEQ ID NO: 66)	GCTGTGTGCCAAGCCTTTC (SEQ ID NO: 138)	TGGCTCCACCACTGATATCCT CCTTTGG (SEQ ID NO: 210)
IL12A	854	AAGATGTACCAGGTGGAGT TCAAGA (SEQ ID NO: 67)	GCCAGCATGTTTGTATCTAG AAAGAT (SEQ ID NO: 139)	CCATGAATGCAAAGCTTCTGA TGGATCCT (SEQ ID NO: 211)
IL6R	856	GCGGCCTGAGGCACGT (SEQ ID NO: 68)	TCGCCTTGCCCGAACTC (SEQ ID NO: 140)	TCCTGGGCACGAAGCTGCACC (SEQ ID NO: 212)
TNFSF8	863	TTCTGTTTTGAGAGCCTCC AGTTAA (SEQ ID NO: 69)	TCATTTCCCATTAGGGCAGA GT (SEQ ID NO: 141)	TGCTAGCTGCTCTGGTACTGG AGCCC (SEQ ID NO: 213)
TNFRSF8	910	GGAGAACTAAGGCTGAAA CCTCG (SEQ ID NO: 70)	CTAGTCCGCACCCAACGC (SEQ ID NO: 142)	AGGAACAACCACTTTTGAAGT GACTTCGCG (SEQ ID NO: 214)
CD38	871	TCATCATACATGACTCAGC ATACCTG (SEQ ID NO: 71)	GTCTCTGGCATTGACCTTAT TGTG (SEQ ID NO: 143)	AGGACCCTCCAAAATCTTCAG CTCTGCA (SEQ ID NO: 215)
PDGFRA	894	CACCTACAAAAACGAGGA AGACAA (SEQ ID NO: 72)	AGCGCTCAGTCTCTGCTCAT C (SEQ ID NO: 144)	CAGACCACCCTCCAGTCCTT CAGC (SEQ ID NO: 216)

D) Analyze RT-PCR data by statistical methods to identify genes differentially expressed (up-regulated) in cancer:

[00453] The expression level of a target gene in both normal and tumor samples was determined using Quantitative RT-PCR using the ABI PRISM 7900HT Sequence Detection System (Applied Biosystems, California). The method is based on the quantitation of the initial copy number of target template in comparison to that of a reference (normalizer) housekeeper gene (Pre-Developed TaqMan® Assay Reagents Gene Expression Quantification Protocol, Applied Biosystems, 2001). Accumulation of DNA product with each PCR cycle is related to amplicon efficiency and the initial template concentration. Therefore the amplification efficiency of both the target and the normalizer must be similar. The threshold cycle (C_T), which is dependent on the starting template copy number and the DNA amplification efficiency, is a PCR cycle during which PCR product growth is exponential. Each assay was performed in quadruplicates; therefore, 4 C_T values were obtained for the target gene in a given sample. Simultaneously, the expression level of a group of housekeeper genes were also measured in the same fashion. The outlier within the 4 quadruplicates is detected and removed if the standard deviation of the remaining 3 triplicates is 30% or less compared to the standard deviation of the original 4 quadruplicates. The mean of the remaining C_T values (designated as C_i or C_n) was calculated and used in the following computation.

Data Normalization.

[00454] For normalization, a 'universal normalizer' was developed that is based on the set of housekeepers available for analysis (5 to 8 genes). Briefly, the housekeeper genes were weighted according to their variations in expression level across the whole panel of tissue samples. For n samples of the same tissue type, the weight (w) for the k th house keeper gene was calculated with the following formulas:

Equation 1

$$w_k = \frac{1/S_k^2}{\sum_{k=1}^n 1/S_k^2}$$

[00455] Where S_k stands for the standard deviation of the k th housekeeper gene across the all samples of same tissue type in the panel. The mean expression of all housekeeper genes in the i th sample (M_i) was estimated using the weighted least square method, and the difference between the M_i and the average of all M_i is computed as the normalization factor N_i for the i th sample (Equation 2). The mean C_t value of the target gene in the i th sample was then normalized by

subtracting the normalization factor N_i . The performance of the above normalization method was validated by comparing the correlation between RT-PCR and microarray data that were generated from the same set of samples: increased correlation between RT-PCR data and microarray data was observed after applying the above normalization method.

Equation 2
$$N_i = M_i - \frac{\sum_{i=1}^n M_i}{n}$$

Identification of significantly dysregulated Genes.

[00456] To determine if a gene is significantly up-regulated in the tumor versus normal samples, two statistics, t (Equation 3) and Receiver Operating Characteristic (ROC; Equation 4) were calculated:

Equation 3
$$t = \frac{\bar{C}_t - \bar{C}_n}{\sqrt{\frac{S_t^2}{n_t} + \frac{S_n^2}{n_n}}}$$

[00457] Equation 4
$$ROC(t_0) = P[C_t \leq C_n(t_0)]$$

where \bar{C}_t is the average of C_t in the tumor sample group, \bar{C}_n is the average of C_n in the normal sample group, S_t , S_n are standard deviations of the tumor and normal control groups, and n_t , n_n are the number of the tumor and normal samples used in the analysis. The degree of freedom v' of t is calculated as:

Equation 5
$$v' = \frac{\left(\frac{S_t^2}{n_t} + \frac{S_n^2}{n_n}\right)^2}{\frac{\left(\frac{S_t^2}{n_t}\right)^2}{n_t - 1} + \frac{\left(\frac{S_n^2}{n_n}\right)^2}{n_n - 1}}$$

[00458] In the ROC equation, t_0 is the accepted false positive rate in the normal population, which is set to 0.1 in our study. Therefore, $C_n(t_0)$ is the 10 percentile of C_n in the normal samples, and the $ROC(0.1)$ is the percentage of tumor samples with C_t lower than the 10

percentile of the normal samples. The t statistic identifies genes that show higher average expression level in tumor samples compared to normal samples, while the ROC statistic is more suitable to identify genes that show elevated expression level only in a subset of tumors. The rationale of using ROC statistic is discussed in detail in Pepe, *et al* (2003) Biometrics 59, 133-142. The distribution of t under null hypothesis is empirically estimated by permutation to avoid normal distribution assumption, in which we randomly assign normal or tumor labels to the samples, and then calculate the t statistic (t^p) as above for 2000 times. The p value was then calculated as the number of t^p less than t from real samples divided by 2000. To assess the variability of ROC , the samples were bootstrapped 2000 times, each time, a bootstrap ROC (ROC^b) was calculated as above. If 97.5% of 2000 ROC^b is above 0.1, the acceptable false positive rate we set for normal population, the ROC from the real samples was then considered as statistically significant. The threshold to determine significance was set at >20% incidence for ROC and <0.05 for the T-test P value.

[00459] Application of the above methodologies allowed us to model 3 hypothetical distributions between the normal and sample sets (Figure 2).

[00460] In scenario I, there was essentially complete separation between the two sample populations (control and disease). Both the ROC and T-Test score this scenario with high significance. In scenario II, the samples exhibit overlapping distributions and only a subset of the disease sample is distinct from the control (normal) population. Only the ROC method will score this scenario as significant. In scenario III, the disease sample population overlaps entirely with the control population. In contrast to scenario I and II, only the T-Test method will score this scenario as significant. In sum, the combination of both statistical methods allows one to accurately characterize the expression pattern of a target gene within a sample population.

Results of these tests are expressed in Table 3.

Table 3: ROC % incidence is given where calculated. Where incidence was not significant the "t" value is given. ns=not significant but no t-value given. Number of insertions and type of insertions are also given along with the virus type used.

Cancer Type, % Incidence vs. Corresponding Normal																
Gene	Insertion Type	No of insert	Virus type	Breast	Colon	Kidney	Liver	Lung	Lymphoid	Ovary	Pancreas	Prostate	Uterine	Cervix	Stomach	Skin
WNT3	I/II	72	M						60%			32%				
NOTCH1	I/II	132	L / M		62%								39%			
WNT1	I/II	346	L/M							57%		t=0.02				
WNT3A	I/II	106	L / M				47%									
IL2RA	I/II	12	L/M		32%	t=0.04			28%			t=0.02	t=0.02			t=0.01
FLT3	I/II	18	L			36%										
TNFRSF6	II	7	L						50%			28%	32%	27%		
PTPRK	II	7	L					33%								
TNFSF11	I	2	L													
FGFR3	I/II	9	L / M	48%						86%						
SELL	II	7	L	38%	t=0.005							t=0.00001		23%		

[illegible]

FCMD	II	2	L			t=0.02							33%		30%		
COL15A1	I/II	7	L/M					50%	38%					29%			
SLAMF6	II	4	L							43%	50%				t=0.04	30%	
CD74	I/II	36	L							23%	t=0.02			26%			
TSPAN-2	I	2	L	45%					26%					t=0.03			
THSD2	I/II	20	M								43%						
CD2	I/II	5	L/M	29%						24%							
LNPEP	I/II	5	L		44%				42%	40%							
IGSF8	II	3	L	52%			55%				65%	30%	41%			30%	
ATP1A2	II	2	L														
ITGA4	I/II	6	L/M	40%													
B2M	I/II	6	L														
PAG	I/II	6	L	t=0.02						33%				40%		43%	
NOTCH2	I/II	40	L / M														
UTS2	I	2	L			t=0.04					37%						
TNFRSF14	I/II	4	L				48%						36%			75%	
TGFB3	I/II	9	L/M										67%	27%			

FGF8	I/II	62	L/M		t=0.0045					33%							
FGF3	I/II	824	L/M														
XYLT1	I/II	4	L/M			31%				40%							
FGF4	I/II	49	M														
LOC390243	I/II	2	L														
THY1	I	4	L	50%	59%		45%	34%	26%		t=-0.05,	54%					
SLC9A9	I/II	4	L			38%	33%										
PRG1	I/II	2	L			t=0.02											
CPM	I/II	4	L/M										t=-0.04				
FBN2	I/II	2	L	48%							29%						
LTB	II	2	L	29%					50%	27%							
KCNH5	I/II	2	M						44%		60%				46%	36%	
SLC12A7	II	4	L/M				48%		t=0.01			50%					
FGF10	I/II	42	M				47%		t=0.04		t=0.03						
IL7R	I/II	7	L					t=0.04	t=0.045								
MATN2	II	2	L/M									32%	24%				
COL19A1	II	3	L/M	27%									24%				

SLC1A2	I/II	2	L							t=0.08			33%	27%	30%
SLC28A2	I	2	L								45%				
SLP1	I/II	5	L							100%					
IL12A	I/II	5	L						47%	t=0.02					
IL6R	I/II	4	L							t=0.01					
TNFSF8	I	2	L												
TNFRSF8	I/II	5	L					56%	33%			39%	45%		44%
CD38	II	3	L / M						30%	35%			t=0.04		
PDGFRA	I/II	11	M						27%	t=0.01					
PRLR	I/II	16	L / M	40%					34%			t=0.03			45%

Example 3: Detection of cancer-associated-Sequences in Human Cancer Cells and Tissues.

[00461] DNA from prostate and breast cancer tissues and other human cancer tissues, human colon, normal human tissues including non-cancerous prostate, and from other human cell lines are extracted following the procedure of Delli Bovi et al. (1986, Cancer Res. 46:6333-6338). The DNA is resuspended in a solution containing 0.05 M Tris HCl buffer, pH 7.8, and 0.1 mM EDTA, and the amount of DNA recovered is determined by microfluorometry using Hoechst 33258 dye. Cesarone, C. et al., Anal Biochem 100:188-197 (1979).

[00462] Polymerase chain reaction (PCR) is performed using Taq polymerase following the conditions recommended by the manufacturer (Perkin Elmer Cetus) with regard to buffer, Mg²⁺, and nucleotide concentrations. Thermocycling is performed in a DNA cycler by denaturation at 94°C. for 3 min. followed by either 35 or 50 cycles of 94°C. for 1.5 min., 50°C. for 2 min. and 72°C. for 3 min. The ability of the PCR to amplify the selected regions of the cancer-associated gene is tested by using a cloned cancer-associated polynucleotide(s) as a positive template(s). Optimal Mg²⁺, primer concentrations and requirements for the different cycling temperatures are determined with these templates. The master mix recommended by the manufacturer is used. To detect possible contamination of the master mix components, reactions without template are routinely tested.

[00463] Southern blotting and hybridization are performed as described by Southern, E. M., (J. Mol. Biol. 98:503-517, 1975), using the cloned sequences labeled by the random primer procedure (Feinberg, A. P., et al., 1983, Anal. Biochem. 132:6-13). Prehybridization and hybridization are performed in a solution containing 6xSSPE, 5% Denhardt's, 0.5% SDS, 50% formamide, 100 µg/ml denaturated salmon testis DNA, incubated for 18 hrs at 42°C., followed by washings with 2xSSC and 0.5% SDS at room temperature and at 37°C. and finally in 0.1xSSC with 0.5% SDS at 68° C. for 30 min (Sambrook et al., 1989, in "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Lab. Press). For paraffin-embedded tissue sections the conditions described by Wright and Manos (1990, in "PCR Protocols", Innis et al., eds., Academic Press, pp. 153-158) are followed using primers designed to detect a 250 bp sequence.

Example 4: Expression of cloned polynucleotides in host cells.

[00464] To study the protein products of cancer-associated genes, restriction fragments from cancer-associated DNA are cloned into the expression vector pMT2 (Sambrook, et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press pp 16.17-

16.22 (1989)) and transfected into COS cells grown in DMEM supplemented with 10% FCS. Transfections are performed employing calcium phosphate techniques (Sambrook, et al (1989) pp. 16.32-16.40, *supra*) and cell lysates are prepared forty-eight hours after transfection from both transfected and untransfected COS cells. Lysates are subjected to analysis by immunoblotting using anti-peptide antibody.

[00465] In immunoblotting experiments, preparation of cell lysates and electrophoresis are performed according to standard procedures. Protein concentration is determined using BioRad protein assay solutions. After semi-dry electrophoretic transfer to nitrocellulose, the membranes are blocked in 500 mM NaCl, 20 mM Tris, pH 7.5, 0.05% Tween-20 (TTBS) with 5% dry milk. After washing in TTBS and incubation with secondary antibodies (Amersham), enhanced chemiluminescence (ECL) protocols (Amersham) are performed as described by the manufacturer to facilitate detection.

Example 5: Generation of antibodies against polypeptides.

[00466] Polypeptides, unique to cancer-associated genes are synthesized or isolated from bacterial or other (*e.g.*, yeast, baculovirus) expression systems and conjugated to rabbit serum albumin (RSA) with m-maleimido benzoic acid N-hydroxysuccinimide ester (MBS) (Pierce, Rockford, Ill.). Immunization protocols with these peptides are performed according to standard methods. Initially, a pre-bleed of the rabbits is performed prior to immunization. The first immunization includes Freund's complete adjuvant and 500µg conjugated peptide or 100µg purified peptide. All subsequent immunizations, performed four weeks after the previous injection, include Freund's incomplete adjuvant with the same amount of protein. Bleeds are conducted seven to ten days after the immunizations.

[00467] For affinity purification of the antibodies, the corresponding cancer-associated polypeptide is conjugated to RSA with MBS, and coupled to CNBr-activated Sepharose (Pharmacia, Uppsala, Sweden). Antiserum is diluted 10-fold in 10 mM Tris-HCl, pH 7.5, and incubated overnight with the affinity matrix. After washing, bound antibodies are eluted from the resin with 100 mM glycine, pH 2.5.

Example 6: Generation of monoclonal antibodies against a cancer-associated polypeptide

[00468] A non-denaturing adjuvant (Ribi, R730, Corixa, Hamilton MT) is rehydrated to 4ml in phosphate buffered saline. 100µl of this rehydrated adjuvant is then diluted with 400µl of Hank's Balanced Salt Solution and this is then gently mixed with the cell pellet used for

immunization. Approximately 500 µg conjugated peptide or 100 µg purified peptide and Freund's complete are injected into Balb/c mice via foot-pad, once a week. After 6 weeks of weekly injection, a drop of blood is drawn from the tail of each immunized animal to test the titer of antibodies against cancer-associated polypeptides using FACS analysis. When the titer reaches at least 1:2000, the mice are sacrificed in a CO₂ chamber followed by cervical dislocation. Lymph nodes are harvested for hybridoma preparation. Lymphocytes from mice with the highest titer are fused with the mouse myeloma line X63-Ag8.653 using 35% polyethylene glycol 4000. On day 10 following the fusion, the hybridoma supernatants are screened for the presence of CAP-specific monoclonal antibodies by fluorescence activated cell sorting (FACS). Conditioned medium from each hybridoma is incubated for 30 minutes with a combined aliquot of PC3, Colo-205, LnCap, or Panc-1 cells. After incubation, the cell samples are washed, resuspended in 0.1 ml diluent and incubated with 1 µg/ml of FITC conjugated F(ab')₂ fragment of goat anti-mouse IgG for 30 min at 4°C. The cells are washed, resuspended in 0.5 ml FACS diluent and analyzed using a FACScan cell analyzer (Becton Dickinson; San Jose, CA). Hybridoma clones are selected for further expansion, cloning, and characterization based on their binding to the surface of one or more of cell lines which express the cancer-associated polypeptide as assessed by FACS. A hybridoma making a monoclonal antibody designated mAbcancer-associated which binds an antigen designated Ag-CA.x and an epitope on that antigen designated Ag-CA.x.1 is selected.

Example 7: ELISA assay for Detecting cancer-associated antigen related antigens.

[00469] To test blood samples for antibodies that bind specifically to recombinantly produced cancer-associated antigens, the following procedure is employed. After a recombinant cancer-associated related protein is purified, the recombinant protein is diluted in PBS to a concentration of 5µg/ml (500 ng/100 µl). 100 microliters of the diluted antigen solution is added to each well of a 96-well Immulon 1 plate (Dynatech Laboratories, Chantilly, Va.), and the plate is then incubated for 1 hour at room temperature, or overnight at 4°C., and washed 3 times with 0.05% Tween 20 in PBS. Blocking to reduce nonspecific binding of antibodies is accomplished by adding to each well 200 µl of a 1% solution of bovine serum albumin in PBS/Tween 20 and incubation for 1 hour. After aspiration of the blocking solution, 100 µl of the primary antibody solution (anticoagulated whole blood, plasma, or serum), diluted in the range of 1/16 to 1/2048 in blocking solution, is added and incubated for 1 hour at room temperature or overnight at 4° C. The wells are then washed 3 times, and 100 µl of goat anti-human IgG antibody conjugated to horseradish peroxidase

(Organon Teknika, Durham, N.C.), diluted 1/500 or 1/1000 in PBS/Tween 20, 100 μ l of *o*-phenylenediamine dihydrochloride (OPD, Sigma) solution is added to each well and incubated for 5-15 minutes. The OPD solution is prepared by dissolving a 5 mg OPD tablet in 50 ml 1% methanol in H₂O and adding 50 μ l 30% H₂O₂ immediately before use. The reaction is stopped by adding 25 μ l of 4M H₂SO₄. Absorbances are read at 490 nm in a microplate reader (Bio-Rad).

Example 8: Identification and characterization of cancer-associated antigen on cancer cell surface

[00470] A cell pellet of proximately 25 μ l packed cell volume of a cancer cell preparation is lysed by first diluting the cells to 0.5 ml in water followed by freezing and thawing three times. The solution is centrifuged at 14,000 rpm. The resulting pellet, containing the cell membrane fragments, is resuspended in 50 μ l of SDS sample buffer (Invitrogen, Carlsbad, CA). The sample is heated at 80°C for 5 minutes and then centrifuged for 2 minutes at 14,000 rpm to remove any insoluble materials.

[00471] The samples are analyzed by Western blot using a 4 to 20% polyacrylamide gradient gel in Tris-Glycine SDS (Invitrogen; Carlsbad CA) following the manufacturer's directions. Ten microliters of membrane sample are applied to one lane on the polyacrylamide gel. A separate 10 μ L sample is reduced first by the addition of 2 μ L of dithiothreitol (100 mM) with heating at 80°C for 2 minutes and then loaded into another lane. Pre-stained molecular weight markers SeeBlue Plus2 (Invitrogen; Carlsbad, CA) are used to assess molecular weight on the gel. The gel proteins are transferred to a nitrocellulose membrane using a transfer buffer of 14.4 g/l glycine, 3 g/l of Tris Base, 10% methanol, and 0.05% SDS. The membranes are blocked, probed with a CAP-specific monoclonal antibody (at a concentration of 0.5 μ g/ml), and developed using the Invitrogen WesternBreeze Chromogenic Kit-AntiMouse according to the manufacturer's directions. In the reduced sample of the tumor cell membrane samples, a prominent band is observed migrating at a molecular weight within about 10% of the predicted molecular weight of the corresponding cancer-associated protein.

Example 9: Preparation of vaccines.

[00472] The present invention also relates to a method of stimulating an immune response against cells that express cancer-associated polypeptides in a patient using cancer-associated polypeptides of the invention that act as an antigen produced by or associated with a malignant cell. This aspect of the invention provides a method of stimulating an immune

response in a human against cancer cells or cells that express cancer-associated polynucleotides and polypeptides. The method comprises the step of administering to a human an immunogenic amount of a polypeptide comprising: (a) the amino acid sequence of a human cancer-associated protein or (b) a mutein or variant of a polypeptide comprising the amino acid sequence of a human endogenous retrovirus cancer-associated protein.

Example 10: Generation of transgenic animals expressing polypeptides as a means for testing therapeutics.

[00473] Cancer-associated nucleic acids are used to generate genetically modified non-human animals, or site specific gene modifications thereof, in cell lines, for the study of function or regulation of prostate tumor-related genes, or to create animal models of diseases, including prostate cancer. The term "transgenic" is intended to encompass genetically modified animals having an exogenous cancer-associated gene(s) that is stably transmitted in the host cells where the gene(s) may be altered in sequence to produce a modified protein, or having an exogenous cancer-associated LTR promoter operably linked to a reporter gene. Transgenic animals may be made through a nucleic acid construct randomly integrated into the genome. Vectors for stable integration include plasmids, retroviruses and other animal viruses, YACs, and the like. Of interest are transgenic mammals, *e.g.* cows, pigs, goats, horses, *etc.*, and particularly rodents, *e.g.* rats, mice, *etc.*

[00474] The modified cells or animals are useful in the study of cancer-associated gene function and regulation. For example, a series of small deletions and/or substitutions may be made in the cancer-associated genes to determine the role of different genes in tumorigenesis. Specific constructs of interest include, but are not limited to, antisense constructs to block cancer-associated gene expression, expression of dominant negative cancer-associated gene mutations, and over-expression of a cancer-associated gene. Expression of a cancer-associated gene or variants thereof in cells or tissues where it is not normally expressed or at abnormal times of development is provided. In addition, by providing expression of proteins derived from cancer-associated in cells in which it is otherwise not normally produced, changes in cellular behavior can be induced.

[00475] DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and negative selection are included. For various techniques for transfecting mammalian cells, see Keown et al., *Methods in Enzymology* 185:527-537 (1990).

[00476] For embryonic stem (ES) cells, an ES cell line is employed, or embryonic cells are obtained freshly from a host, e.g. mouse, rat, guinea pig, etc. Such cells are grown on an appropriate fibroblast-feeder layer or grown in the presence of appropriate growth factors, such as leukemia inhibiting factor (LIF). When ES cells are transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting chimeric animals screened for cells bearing the construct. By providing for a different phenotype of the blastocyst and the ES cells, chimeric progeny can be readily detected.

[00477] The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs are maintained as allogeneic or congenic grafts or transplants, or in in vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animals, domestic animals, etc. The transgenic animals are used in functional studies, drug screening, etc., e.g. to determine the effect of a candidate drug on prostate cancer, to test potential therapeutics or treatment regimens, etc.

Example 11: Diagnostic Imaging Using CA Specific Antibodies

[00478] The present invention encompasses the use of antibodies to cancer-associated polypeptides to accurately stage cancer patients at initial presentation and for early detection of metastatic spread of cancer. Radioimmunoscinigraphy using monoclonal antibodies specific for cancer-associated polypeptides can provide an additional cancer-specific diagnostic test. The monoclonal antibodies of the instant invention are used for histopathological diagnosis of carcinomas.

[00479] Subcutaneous human xenografts of cancer cells in nude mice are used to test whether a technetium-99m (^{99m}Tc)-labeled monoclonal antibody of the invention can

successfully image the xenografted cancer by external gamma scintigraphy as described for seminoma cells by Marks, *et al.*, Brit. J. Urol. 75:225 (1995). Each monoclonal antibody specific for a cancer-associated polypeptide is purified from ascitic fluid of BALB/c mice bearing hybridoma tumors by affinity chromatography on protein A-Sepharose. Purified antibodies, including control monoclonal antibodies such as an avidin-specific monoclonal antibody (Skea, *et al.*, J. Immunol. 151:3557 (1993)) are labeled with ^{99m}Tc following reduction, using the methods of Mather, *et al.*, J. Nucl. Med. 31:692 (1990) and Zhang *et al.*, Nucl. Med. Biol. 19:607 (1992). Nude mice bearing human cancer cells are injected intraperitoneally with 200-500 μCi of ^{99m}Tc -labeled antibody. Twenty-four hours after injection, images of the mice are obtained using a Siemens ZLC3700 gamma camera equipped with a 6 mm pinhole collimator set approximately 8 cm from the animal. To determine monoclonal antibody biodistribution following imaging, the normal organs and tumors are removed, weighed, and the radioactivity of the tissues and a sample of the injectate are measured. Additionally, cancer-associated antigen-specific antibodies conjugated to antitumor compounds are used for cancer-specific chemotherapy.

Example 12: Immunohistochemical methods

[00480] Frozen tissue samples from cancer patients are embedded in an optimum cutting temperature (OCT) compound and quick-frozen in isopentane with dry ice. Cryosections are cut with a Leica 3050 CM microtome at thickness of 5 μm and thaw-mounted on vectabound-coated slides. The sections are fixed with ethanol at -20°C and allowed to air dry overnight at room temperature. The fixed sections are stored at -80°C until use. For immunohistochemistry, the tissue sections are retrieved and first incubated in blocking buffer (PBS, 5% normal goat serum, 0.1% Tween 20) for 30 minutes at room temperature, and then incubated with the cancer-associated protein-specific monoclonal antibody and control monoclonal antibodies diluted in blocking buffer (1 $\mu\text{g}/\text{ml}$) for 120 minutes. The sections are then washed three times with the blocking buffer. The bound monoclonal antibodies are detected with a goat anti-mouse IgG + IgM (H+L) F(ab') $_2$ -peroxidase conjugates and the peroxidase substrate diaminobenzidine (1 mg/ml, Sigma Catalog No. D 5637) in 0.1 M sodium acetate buffer pH 5.05 and 0.003% hydrogen peroxide (Sigma cat. No. H1009). The stained slides are counter-stained with hematoxylin and examined under Nikon microscope.

[00481] Monoclonal antibody against a cancer-associated protein (antigen) is used to test reactivity with various cell lines from different types of tissues. Cells from different established cell lines are removed from the growth surface without using proteases, packed

and embedded in OCT compound. The cells are frozen and sectioned, then stained using a standard IHC protocol. The CellArray™ technology is described in WO 01/43869. Normal tissue (human) obtained by surgical resection are frozen and mounted. Cryosections are cut with a Leica 3050 CM microtome at thickness of 5 µm and thaw-mounted on vectabound-coated slides. The sections are fixed with ethanol at -20°C and allowed to air dry overnight at room temperature. PolyMICA™ Detection kit is used to determine binding of a cancer-associated antigen-specific monoclonal antibody to normal tissue. Primary monoclonal antibody is used at a final concentration of 1 µg/ml.

Example 13: Expression Data

[00482] Tables 4-58 show the expression level of a number of genes from this application in a range of tumor tissues. The results of three forms of expression assay are shown in the table. Results with the notation U133A&B (see Array Type column) indicate that an Affymetrix oligonucleotide based expression array was used (worldwide web site: affymetrix.com/support/technical/byproduct.affx?product=hg-u133-plus) results with the notation Chiron cDNA array were produced using a Chiron in house spotted cDNA array. The remaining results were produced using the Affymetrix U133 plus 2 oligo array (worldwide web site: affymetrix.com/support/technical/byproduct.affx?product=hg-u133-plus).

[00483] Tissue samples for the first two array formats were collected using laser capture microdissection (LCD) (see definition below). Tissue samples for the third form of array (U133 plus 2) were collected using standard manual dissection procedures.

[00484] The level of differential expression was assessed by separate methods for each array type. The results for U133A&B and EVD arrays are expressed as the number of samples that had an expression level either above or below a defined threshold. The notation “2X concordance” or “0.5X Concordance” indicates that a number of the samples (shown in the column % samples showing diff level of expression) had either a 2X greater level of expression than the control values or a level of expression less than half that of the control average. Thus showing either over or under-expression of the gene (at a significance level of “t” ≤ 0.001).

[00485] The expression levels for the U133+2 arrays are expressed as 80-50 or 80-80 percentile ratios. All results shown have passed a t test for significance at the 0.001% level. In an 80-50 Percentile Ratio result 20% of the tumor samples have a greater the 2 fold level of

overexpression when compared to 50% of the control samples. In an 80-80 percentile ratio result 20% of the tumor samples have a greater than 3 fold level of overexpression as compared to 80% of the control samples.

Selection of Tumor Associated Antigens for targeting

Laser dissection of tumorous cells and adjacent normals and production of RNA from dissected cells.

[00486] Normal and cancerous tissues were collected from patients using laser capture microdissection (LCM), and RNA was prepared from these tissues, using techniques which are well known in the art (see, e.g., Ohyama et al. (2000) *Biotechniques* 29:530-6; Curran et al. (2000) *Mol. Pathol.* 53:64-8; Suarez-Quian et al. (1999) *Biotechniques* 26:328-35; Simone et al. (1998) *Trends Geriatr* 14:272-6; Conia et al. (1997) *J. Clin. Lab. Anal.* 11:28-38; Emmert-Buck et al. (1996) *Science* 274:998-1001). Because LCM provides for the isolation of specific cell types to provide a substantially homogenous cell sample, this provided for a similarly pure RNA sample.

Microarray analysis

[00487] Production of cDNA: Total RNA produced from the dissected cells was then used to produce cDNA using an Affymetrix Two-cycle cDNA Synthesis Kit (cat# 900432). 8 μ L of total RNA was used with 1 μ L T7-(dT) 24 primer (50 pmol/ μ L) in an 11 μ L reaction which was heated to 70° C for 12 minutes. The mixture was then cooled to room temperature for five minutes. 9 μ L master mix (4 μ L 5x 1st strand cDNA buffer, 2 μ L 0.1 M DTT, 1 μ L 10 mM dNTP mix, 2 μ L Superscript II (600 U/ μ L)) was added and the mixture was incubated for 2.5 hours at 42° C (total volume of the mixture was 20 μ L). Following cooling on ice, the 2nd strand synthesis was completed as follows: 20 μ L mixture from above was mixed with 130 μ L second strand master mix (91 μ L water, 30 μ L 5x Second Strand Reaction Buffer, 3 μ L 10 mM dNTP mix, 1 μ L 10 U/ μ L e. coli DNA ligase, 4 μ L 10 U/ μ L E. coli DNA polymerase I, 1 μ L 2 U/ μ L e. coli Rnase H) and was incubated for 2 hours at 16° C for 10 minutes.

Following cooling on ice, the dsDNA was purified from the reaction mixture. Briefly, a QiaQuick PCT Purification Kit was used (Qiagen, cat# 28104), and 5 volumes of buffer PB was added to 1 volume of the cDNA mixture. The cDNA was then purified on a QIAquick spin column according to manufacture's directions, yielding a final volume of 60 μ L.

[00488] Production of biotin-labeled cRNA. The cDNA produced and purified above was then used to make biotin labeled RNA as follows: The 60 μ L of cDNA recovered from the

QIAQuick column was reduced to a volume of 22 μ L in a medium heated speed vacuum. This was then used with an ENZO BioArray High Yield RNA Transcription Kit (cat# 4265520). Briefly, a master mix containing 4 μ L 10x HY Reaction buffer, 4 μ L 10x Biotin-Labeled Ribonucleotides, 4 μ L DTT, 4 μ L Rnase Inhibitor Mix, and 2 μ L T7 RNA Polymerase was added to the 22 μ L of purified cDNA, and left to incubate at 37 °C for 4 to 6 hours. The reaction was then purified using a Qiagen RNeasy Kit (cat# 74104) according to manufacturer's directions.

[00489] Fragmentation of cRNA. 15 to 20 μ g of cRNA from above was mixed with 8 μ L of 5x Fragmentation Buffer (200mM Tris-acetate, pH 8.1, 500mM Potassium acetate, 150mM Magnesium acetate) and water to a final volume of 40 μ L. The mixture was incubated at 94 °C for 35 minutes. Typically, this fragmentation protocol yields a distribution of RNA fragments that range in size from 35 to 200 bases. Fragmentation was confirmed using TAE agarose electrophoresis.

[00490] Array Hybridization. The fragmented cRNA from above was then used to make a hybridization cocktail. Briefly, the 40 μ L from above was mixed with 1 mg/mL human Cot DNA and a suitable control oligonucleotide. Additionally, 3 mg of Herring Sperm DNA (10 mg/mL) was added along with 150 μ L 2x Hybridization buffer (100 mM MES, 1 M NaCl, 20 mM EDTA, 0.01 % Tween-20) and water to a final volume of 300 μ L. 200 μ L of this solution was then loaded onto the U133 array (Affymetrix cat # 900370) and incubated at 45 °C with a constant speed of 45 rpm overnight. The hybridization buffer was then removed and the array was washed and stained with 200 μ L Non-stringent wash buffer (6X SSPE, 0.01% Tween-20) and using a GeneChip Fluidics Station 450 (Affymetrix, cat# 00-0079) according to manufacturer's protocol.

[00491] Scanning array. The array from above was then scanned using a GeneChip Scanner 3000 (Affymetrix cat# 00-0217) according to manufacturer's protocol.

[00492] Selection of potential tumor cell antigen targets. The tumor antigens were selected for targeting by comparison of the expression level of the antigen in the tumor cells (either primary tumors or metastases) versus neighboring healthy tissue or with pooled normal tissue. Tumor antigens selected showed at least a 3 fold (300%) increased expression relative to surrounding normal tissue, where this 3 fold increase is seen in comparison with a majority of pooled, commercially available normal tissue samples (Reference standard mix or RSM, pools are made for each tissue type). The tables below present the fold increase data from the

array analysis for the respective genes, where the numbers represent the percent of patient samples analyzed that showed a 2-, 3- or 5-fold increase in expression in comparison to normal tissues.

Table 4

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
ANXA2	Chiron cDNA array	Breast Cancer	2X Concordance	23.0		26.0		1.0
ANXA2	Chiron cDNA array	Colon Cancer	2X Concordance	68.0	6.9	22.0	1.0	3.0
ANXA2	Chiron cDNA array	matched Colon Cancer	2X Concordance	27.0	6.0	30.5	3.8	4.0
ANXA2	Chiron cDNA array	matched Colon Cancer metastasis	2X Concordance	26.0	6.0	30.0	3.3	4.0
ANXA2	Chiron cDNA array	Prostate Cancer	<=0.5X Concordance	89.2	30.9	54.6	12.2	5.0
ANXA2	U133A & B	Colon Cancer	2X Concordance	27.0	0.0	30.7	4	3.0
ANXA2	U133A & B	Colon metastasis	2X Concordance	33.0	0.0	24.0	4.2	2.0
ANXA2	U133A & B	Prostate Cancer	<=0.5X Concordance	22.0	0.0	89.3	2.9	3.0
ANXA2	U133A & B	Prostate	<=0.5X	15.7	0.6	68.0	6.6	3.0

	B	Cancer vs adjacent Normal	Concordance	Avg No of Tissue Samples	SDev of Sample No	Fold Diff from normal	SDev of Fold or % diff.	No of Replicate Experiments
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>					
ANXA2	U133A & B	Bladder Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	20.0	0.0	8.0	1.0	3.0
ANXA2	U133A & B	Breast Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	43.0	0.0	5.0	0.0	3.0
ANXA2	U133A & B	Lung Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	68.0	0.0	7.3	0.6	3.0
ANXA2	U133A & B	Stomach Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	23.0	0.0	7.0	0.0	3.0
ANXA2	U133A & B	Upper Aero-Digestive Tract Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	23.0	0.0	8.7	0.6	3.0
ANXA2	U133A & B	Uterus Carcinoma All Vs Essential	80-50 Percentile Ratio	39.0	0.0	6.0	0.0	3.0

[illegible]

Table 5

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
ATP1A2	Chiron cDNA array	Colon metastasis	<=0.5X Concordance	33.0	30.0	1.0
ATP1A2	U133A & B	Prostate Cancer	2X Concordance	5.0	100.0	1.0

Table 6

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
B2M	Chiron cDNA array	Breast Cancer	$\leq 0.5X$ Concordance	23.0	0.0	22.0	0.0	3.0
B2M	Chiron cDNA array	Breast Cancer	2X Concordance	23.0	0.0	31.9	4.2	25.0
B2M	Chiron cDNA	Colon Cancer	$\leq 0.5X$ Concordance	67.8	19.2	45.8	3.3	27.0

B2M	array									
	Chiron cDNA array	Colon metastasis	<=0.5X Concordance	32.9	0.6	67.1	3.4	24.0		
B2M	Chiron cDNA array	matched Colon Cancer Cancer	<=0.5X Concordance	32.9	7.9	56.8	5.9	27.0		
B2M	Chiron cDNA array	matched Colon Cancer metastasis	<=0.5X Concordance	33.2	6.7	42.6	3.4	27.0		
B2M	Chiron cDNA array	Prostate Cancer	<=0.5X Concordance	102.8	13.6	35.1	4.2	28.0		
B2M	U133A & B	Breast Cancer	2X Concordance	35.5	20.5	24.0	0.0	2.0		
B2M	U133A & B	Colon Cancer	<=0.5X Concordance	27.0		52.0		1.0		
B2M	U133A & B	Colon metastasis	<=0.5X Concordance	33.0		45.0		1.0		
B2M	U133A & B	Prostate Cancer	<=0.5X Concordance	19.5	3.5	72.0	5.7	2.0		

B2M	U133A & B	Prostate Cancer vs adjacent Normal	<=0.5X Concordance	16.0	SDev of Sample No	Fold Diff from normal	SDev of Fold or % diff.	1.0
Gene Symbol	Array Type	Tumour Type and Control Notation	Result Type	Avg No of Tissue Samples	SDev of Sample No	Fold Diff from normal	SDev of Fold or % diff.	No of Replicate Experiments
B2M	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0		3.0		1.0
B2M	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	0.0	4.0	0.0	2.0
B2M	U133A & B	Liver Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	21.0		4.0		1.0
B2M	U133A & B	Lung Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	68.0		3.0		1.0
B2M	U133A & B	Stomach Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	23.0		3.0		1.0
B2M	U133A & B	Upper Aero-Digestive Tract Carcinoma All Vs	80-50 Percentile Ratio	23.0		2.0		1.0

[illegible]

Table 7

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
CD2	Chiron cDNA array	Breast Cancer	2X Concordance	23.0	26.0	1.0
CD2	Chiron cDNA array	matched Colon Cancer metastasis	<=0.5X Concordance	18.0	22.0	1.0
					Fold Diff from normal	
CD2	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0	8.0	1.0
CD2	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	9.0	1.0

Table 8

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
--------------------	-------------------	---	--------------------	---------------------------------	---	------------------------------------

CD28	U133A & B	Prostate Cancer	<=0.5X Concordance	22.0			23.0	1.0	
CD28	U133A & B	Prostate Cancer vs adjacent Normal	<=0.5X Concordance	16.0			25.0	1.0	
Gene Symbol	Array Type	Tumour Type and Control Notation	Result Type	Avg No of Tissue Samples	SDev of Sample No	Fold Diff from normal	SDev of Fold or % diff.	No of Replicate Experiments	
CD28	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0	0.0	6.3	2.5	3.0	
CD28	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	0.0	9.3	2.3	3.0	
CD28	U133A & B	Liver Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	21.0		2.0		1.0	
CD28	U133A & B	Lung Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	68.0		2.0		1.0	
CD28	U133A & B	Melanoma All Vs Essential Normal	80-50 Percentile Ratio	41.0	0.0	3.0	0.0	2.0	
CD28	U133A & B	Renal Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	66.0		2.0		1.0	
CD28	U133A & B	Stomach Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	23.0		2.0		1.0	

Table 9

Gene Symbol	Array Type	Tumour Type and Control Notation	Result Type	Avg No of Tissue Samples	% samples showing diff level of expression	No of Replicate Experiments
CD38	Chiron	Colon metastasis	<=0.5X Concordance	33.0	24.0	1.0

	cDNA array						
CD38	Chiron cDNA array	Prostate Cancer	<=0.5X Concordance	106.0	52.0	1.0	
CD38	U133A & B Array Type	Colon Cancer	<=0.5X Concordance	25.0	24.0	1.0	
Gene Symbol		<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>Fold Diff from normal</u>	<u>No of Replicate Experiments</u>	
CD38	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0	6.0	1.0	
CD38	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	11.0	1.0	

Table 10

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
CD74	Chiron cDNA array	Breast Cancer	2X Concordance	22.0	2.2	27.8	5.7	5.0
CD74	Chiron cDNA array	Colon Cancer	<=0.5X Concordance	60.2	24.4	22.3	0.8	6.0
CD74	Chiron cDNA array	Colon Cancer	2X Concordance	60.2	24.4	29.0	6.9	6.0

	array									
CD74	Chiron cDNA array	Colon metastasis	<=0.5X Concordance	32.4	1.3	28.6	6.2	5.0		
CD74	Chiron cDNA array	matched Colon Cancer	2X Concordance	31.5	9.0	27.5	11.0	4.0		
CD74	Chiron cDNA array	matched Colon Cancer Cancer	<=0.5X Concordance	32.4	8	22.6	1.3	5.0		
CD74	Chiron cDNA array	matched Colon Cancer metastasis	<=0.5X Concordance	31.0	9.3	24.0	3.9	4.0		
CD74	Chiron cDNA array	matched Colon Cancer metastasis	2X Concordance	31.8	8.3	26.8	4.9	5.0		
CD74	Chiron cDNA array	Prostate Cancer	<=0.5X Concordance	98.3	16.9	26.5	2.6	6.0		
CD74	U133A & B	Colon Cancer	<=0.5X Concordance	22.0		23.0		1.0		

CD74	U133A & B	Colon metastasis	<=0.5X Concordance	33.0		27.0	1.0
CD74	U133A & B	Prostate Cancer	<=0.5X Concordance	22.0		32.0	1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>Fold Diff from normal</u>	<u>No of Replicate Experiments</u>
CD74	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0		6.0	1.0
CD74	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0		14.0	1.0

Table 11

CDH23	U133A & B	Colon Cancer	<=0.5X Concordance	19.0		32.0	1.0
CDH23	U133A & B	Colon Cancer	2X Concordance	26.0		23.0	1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>

CDH23	U133A & B	Colon metastasis	<=0.5X Concordance	32.5	0.7	29.5	12	2.0
CDH23	U133A & B	Prostate Cancer	2X Concordance	22.0		23.0		1.0

Table 12

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
COL15A1	Chiron cDNA array	Colon Cancer	<=0.5X Concordance	19.0	32.0	1.0
COL15A1	Chiron cDNA array	matched Colon Cancer Cancer	<=0.5X Concordance	18.0	33.0	1.0
COL15A1	Chiron cDNA array	matched Colon Cancer metastasis	<=0.5X Concordance	18.0	33.0	1.0
COL15A1	U133A & B	Colon Cancer	2X Concordance	19.0	32.0	1.0

Table 13

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
COL19A1	U133A & B	Colon metastasis	2X Concordance	10.0	100.0	1.0

Table 14

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
CPM	Chiron cDNA array	Colon Cancer	<=0.5X Concordance	76.5	0.7	46.5	17.7	2.0
CPM	Chiron cDNA array	Colon metastasis	<=0.5X Concordance	33.0	0.0	62.5	6.4	2.0
CPM	Chiron cDNA array	matched Colon Cancer	<=0.5X Concordance	36.0	0.0	43.0	9.9	2.0
CPM	Chiron cDNA array	matched Colon Cancer metastasis	<=0.5X Concordance	36.0	0.0	39.0	11.3	2.0
CPM	Chiron cDNA array	Prostate Cancer	<=0.5X Concordance	106.0	0.0	23.0	1.4	2.0
CPM	U133A & B	Breast Cancer	<=0.5X Concordance	50.0		28.0		1.0
CPM	U133A &	Breast Cancer	2X Concordance	11.0		100.0		1.0

	B								
CPM	U133A & B	Colon Cancer	<=0.5X Concordance	25.3	0.6	70.7	33.5	3.0	
CPM	U133A & B	Colon metastasis	<=0.5X Concordance	28.5	9.0	73.0	25.4	4.0	
CPM	U133A & B	Prostate Cancer vs adjacent Normal	<=0.5X Concordance	15.5	3.0	63.0	19.3	4.0	

Table 15

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
DPT	U133A & B	Breast Cancer	2X Concordance	13.5	10.6	66.5	47.4	2.0

Table 16

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
FAS	U133A & B	Colon Cancer	<=0.5X Concordance	23.5	4.9	30.0	0.0	2.0
FAS	U133A	Colon Cancer	2X Concordance	27.0		44.0		1.0

	& B								
FAS	U133A & B	Colon metastasis	<=0.5X Concordance	29.5	4.9	51.0	1.4	2.0	
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>Fold Diff from normal</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>	
FAS	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0		3.0		1.0	
FAS	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	0.0	4.8	1.0	4.0	
FAS	U133A & B	Renal Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	66.0	0.0	3.8	1.0	4.0	

Table 17

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
FBN2	U133A & B	Prostate Cancer	<=0.5X Concordance	19.0	42.0	1.0

Table 18

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
------------------------	-----------------------	---	--------------------	---	---	--

FGF8	U133A & B	Colon Cancer	<=0.5X Concordance	25.0	52.0	1.0
FGF8	U133A & B	Colon metastasis	<=0.5X Concordance	32.0	34.0	1.0

Table 19

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
FGFR1	Chiron cDNA array	Breast Cancer	<=0.5X Concordance	23.0	0.0	48.0	12.7	2.0
FGFR1	Chiron cDNA array	Colon metastasis	<=0.5X Concordance	33.0		21.0		1.0
FGFR1	U133A & B	Breast Cancer	<=0.5X Concordance	49.5	0.7	39.5	12	2.0
FGFR1	U133A & B	Colon Cancer	<=0.5X Concordance	27.0		41.0		1.0
FGFR1	U133A & B	Colon Cancer	2X Concordance	26.0		62.0		1.0

FGFR1	U133A & B	Colon metastasis	2X Concordance	33.0		61.0		1.0
FGFR1	U133A & B	Prostate Cancer	2X Concordance	20.0		30.0		1.0

Table 20

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
FGFR2	Chiron cDNA array	Breast Cancer	<=0.5X Concordance	23.0	0.0	47.5	24.7	2.0
FGFR2	Chiron cDNA array	Colon Cancer	<=0.5X Concordance	19.0		32.0		1.0
FGFR2	Chiron cDNA array	matched Colon Cancer Cancer	<=0.5X Concordance	18.0		33.0		1.0
FGFR2	Chiron cDNA array	matched Colon Cancer metastasis	<=0.5X Concordance	18.0		50.0		1.0

FGFR2	Chiron cDNA array	Prostate Cancer	<=0.5X Concordance	103.5	3.5	36.5	9.2	2.0
FGFR2	U133A & B	Breast Cancer	<=0.5X Concordance	50.0		46.0		1.0
FGFR2	U133A & B	Colon Cancer	<=0.5X Concordance	24.7	1.5	40.3	29.4	3.0
FGFR2	U133A & B	Colon metastasis	<=0.5X Concordance	31.5	0.7	58.5	46.0	2.0
FGFR2	U133A & B	Prostate Cancer	<=0.5X Concordance	18.0	4.2	57.0	14.1	2.0
FGFR2	U133A & B	Prostate Cancer vs adjacent Normal	<=0.5X Concordance	14.0	2.8	63.5	27.6	2.0

Table 21

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
FGFR3	Chiron cDNA array	Breast Cancer	2X Concordance	23.0		26.0		1.0

FGFR3	Chiron cDNA array	Colon Cancer	<=0.5X Concordance	76.0	0.0	31.5	13.4	2.0
FGFR3	Chiron cDNA array	Colon metastasis	<=0.5X Concordance	33.0		30.0		1.0
FGFR3	Chiron cDNA array	matched Colon Cancer Cancer	<=0.5X Concordance	36.0		44.0		1.0
FGFR3	Chiron cDNA array	matched Colon Cancer metastasis	<=0.5X Concordance	35.0		34.0		1.0
FGFR3	U133A & B	Breast Cancer	2X Concordance	26.0		100.0		1.0
FGFR3	U133A & B	Colon Cancer	<=0.5X Concordance	26.0		54.0		1.0
FGFR3	U133A & B	Colon metastasis	<=0.5X Concordance	32.0		41.0		1.0
FGFR3	U133A & B	Prostate Cancer vs adjacent Normal	<=0.5X Concordance	16.0		25.0		1.0

Table 22

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
IFNAR2	U133A & B	Breast Cancer	2X Concordance	28.0	100.0	1.0
IFNAR2	U133A & B	Colon Cancer	2X Concordance	27.0	33.0	1.0
IFNAR2	U133A & B	Colon metastasis	2X Concordance	31.0	32.0	1.0
IFNAR2	U133A & B	Prostate Cancer	2X Concordance	7.0	100.0	1.0
IFNAR2	U133A & B	Prostate Cancer vs adjacent Normal	<=0.5X Concordance	15.0	20.0	1.0
IFNAR2	U133A & B	Prostate Cancer vs adjacent Normal	2X Concordance	15.0	20.0	1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>Fold Diff from normal</u>	<u>No of Replicate Experiments</u>
IFNAR2	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	2.0	1.0

Table 23

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
IGSF8	Chiron cDNA array	Breast Cancer	2X Concordance	23.0	22.0	1.0
IGSF8	Chiron cDNA array	Colon Cancer	<=0.5X Concordance	76.0	21.0	1.0
IGSF8	Chiron cDNA array	Colon metastasis	<=0.5X Concordance	33.0	33.0	1.0
IGSF8	U133A & B	Colon Cancer	<=0.5X Concordance	21.0	24.0	1.0

Table 24

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
IL12A	U133A & B	Colon Cancer	<=0.5X Concordance	22.0	32.0	1.0

Table 25

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of</u>	<u>SDev of Sample No</u>	<u>SDev of Fold</u>	<u>No of Replicate Experiments</u>
--------------------	-------------------	---	--------------------	---------------------------------	--	--------------------------	---------------------	------------------------------------

							<u>expression</u>	<u>or % diff.</u>	
IL2RA	Chiron cDNA array	Breast Cancer	2X Concordance	23.0			22.0	1.0	
IL2RA	Chiron cDNA array	Colon metastasis	2X Concordance	33.0			21.0	1.0	
IL2RA	Chiron cDNA array	matched Colon Cancer	2X Concordance	36.0			28.0	1.0	
IL2RA	Chiron cDNA array	matched Colon Cancer metastasis	2X Concordance	35.0			20.0	1.0	
IL2RA	U133A & B	Colon Cancer	<=0.5X Concordance	26.0			35.0	1.0	
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>Fold Diff from normal</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>	
IL2RA	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	0.0	4.0	0.0	2.0	

Table 26

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>Fold Diff from normal</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
IL4R	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0		6.0		1.0
IL4R	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0		9.0		1.0

Table 27

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
IL6R	U133A & B	Breast Cancer	<=0.5X Concordance	24.0		21.0		1.0
IL6R	U133A & B	Breast Cancer	2X Concordance	24.0		21.0		1.0
IL6R	U133A & B	Colon Cancer	<=0.5X Concordance	21.7	7.6	77.0	20	3.0
IL6R	U133A & B	Colon metastasis	<=0.5X Concordance	24.0	9.9	66.5	38.9	2.0

IL6R	U133A & B	Prostate Cancer	<=0.5X Concordance	22.0		32.0	1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>Fold Diff from normal</u>	<u>SDev of Fold or % diff.</u> <u>No of Replicate Experiments</u>
IL6R	U133A & B	Liver Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	21.0		6.0	1.0
IL6R	U133A & B	Liver Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	21.0		4.0	1.0

Table 28

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>Fold Diff from normal</u>	<u>SDev of Fold or % diff.</u> <u>No of Replicate Experiments</u>
ITGA4	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0	0.0	7.3	3.0
ITGA4	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	0.0	9.3	3.0
ITGA4	U133A & B	Renal Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	66.0		3.0	1.0

Table 29

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
LFNG	Chiron cDNA array	Breast Cancer	2X Concordance	18.0	22.0	1.0
LFNG	U133A & B	Breast Cancer	2X Concordance	43.0	35.0	1.0
LFNG	U133A & B	Colon Cancer	<=0.5X Concordance	26.0	27.0	1.0
LFNG	U133A & B	Colon Cancer	2X Concordance	26.0	58.0	1.0
LFNG	U133A & B	Colon metastasis	2X Concordance	33.0	27.0	1.0
LFNG	U133A & B	Prostate Cancer	2X Concordance	20.0	30.0	1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>Fold Diff from normal</u>	<u>No of Replicate Experiments</u>
LFNG	U133A & B	Breast Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	43.0	6.0	1.0
LFNG	U133A & B	Breast Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	43.0	4.0	1.0

LFNG	U133A & B	Lung Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	68.0	4.0	1.0
LFNG	U133A & B	Pancreas Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	13.0	4.0	1.0
LFNG	U133A & B	Stomach Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	23.0	5.0	1.0

Table 30

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>Fold Diff from normal</u>	<u>No of Replicate Experiments</u>
LNPEP	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	3.0	1.0

Table 31

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
LTB	Chiron cDNA array	Breast Cancer	2X Concordance	23.0	0.0	28.0	2.8	2.0
LTB	Chiron cDNA	matched Colon Cancer	2X Concordance	36.0		22.0		1.0

	array						
LTB	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0		29.0	1.0
LTB	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0		36.0	1.0

Table 32

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
LTBR	U133A & B	Colon metastasis	2X Concordance	33.0	61.0	1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>Fold Diff from normal</u>	<u>No of Replicate Experiments</u>
LTBR	U133A & B	Bladder Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	20.0	2.0	1.0

Table 33

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
MATN2	Chiron cDNA array	Breast Cancer	<=0.5X Concordance	23.0	0.0	76.0	2.8	2.0

MATN2	Chiron cDNA array	Colon Cancer	<=0.5X Concordance	76.5	0.7	26.0	0.0	2.0
MATN2	Chiron cDNA array	Colon metastasis	<=0.5X Concordance	33.0		27.0		1.0
MATN2	Chiron cDNA array	matched Colon Cancer metastasis	<=0.5X Concordance	36.0	0.0	36.0	11.3	2.0
MATN2	U133A & B	Breast Cancer	<=0.5X Concordance	50.0		88.0		1.0
MATN2	U133A & B	Colon Cancer	<=0.5X Concordance	26.0		50.0		1.0
MATN2	U133A & B	Colon metastasis	<=0.5X Concordance	32.0		75.0		1.0

Table 34

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
MET	Chiron	Breast Cancer	<=0.5X Concordance	23.0		30.0		1.0

[illegible]

	& B								
MET	U133A & B	Prostate Cancer vs adjacent Normal	<=0.5X Concordance	10.0				80.0	1.0
MET	U133A & B	Bladder Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	20.0	0.0			12.0	2.0
MET	U133A & B	Bladder Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	20.0				5.0	1.0
MET	U133A & B	Liver Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	21.0	0.0			21.7	3.0
MET	U133A & B	Liver Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	21.0	0.0			9.0	2.0
MET	U133A & B	Lung Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	68.0	0.0			18.7	3.0
MET	U133A & B	Lung Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	68.0	0.0			8.0	2.0
MET	U133A & B	Melanoma All Vs Essential Normal	80-50 Percentile Ratio	41.0	0.0			33.5	2.0
MET	U133A & B	Melanoma All Vs Essential Normal	80-80 Percentile Ratio	41.0	0.0			10.5	2.0
MET	U133A & B	Ovarian Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	72.0				22.0	1.0

	& B											
MET	U133A & B	Ovarian Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	72.0				5.0				1.0
MET	U133A & B	Pancreas Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	13.0				20.0				1.0
MET	U133A & B	Pancreas Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	13.0				5.0				1.0
MET	U133A & B	Renal Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	66.0			0.0	25.0		37.0		4.0
MET	U133A & B	Renal Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	66.0			0.0	11.0		8.2		3.0
MET	U133A & B	Stomach Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	23.0				19.0				1.0
MET	U133A & B	Upper Aero-Digestive Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	23.0				5.0				1.0
MET	U133A & B	Upper Aero-Digestive Tract Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	23.0				19.0				1.0
MET	U133A & B	Uterus Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	39.0			0.0	11.5		10.6		2.0
MET	U133A	Uterus Endometastasis Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	39.0				5.0				1.0

[illegible]

Table 35

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
NOTCH1	Chiron cDNA array	Colon Cancer	2X Concordance	35.0	20.0	1.0
NOTCH1	Chiron cDNA array	Colon metastasis	2X Concordance	30.0	20.0	1.0
NOTCH1	U133A & B	Colon Cancer	2X Concordance	26.0	38.0	1.0

Table 36

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
NOTCH2	Chiron cDNA array	Breast Cancer	2X Concordance	23.0		39.0		1.0
NOTCH2	U133A & B	Colon Cancer	$\leq 0.5X$ Concordance	27.0		30.0		1.0

NOTCH2	U133A & B	Colon metastasis	<=0.5X Concordance	33.0	0.0	24.0	4.2	2.0
NOTCH2	U133A & B	Colon metastasis	2X Concordance	33.0		21.0		1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>Fold Diff from normal</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
NOTCH2	U133A & B	Breast Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	43.0	0.0	2.0	0.0	2.0

Table 37

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
PAG1	Chiron cDNA array	Colon Cancer	<=0.5X Concordance	41.0		29.0		1.0
PAG1	Chiron cDNA array	Colon metastasis	<=0.5X Concordance	30.0		20.0		1.0
PAG1	U133A & B	Breast Cancer	2X Concordance	27.0		48.0		1.0
PAG1	U133A	Colon Cancer	<=0.5X Concordance	26.0		77.0		1.0

	& B									
PAG1	U133A & B	Colon metastasis	<=0.5X Concordance	33.0					64.0	1.0
PAG1	U133A & B	Prostate Cancer	<=0.5X Concordance	22.0					45.0	1.0
Gene Symbol	Array Type	Tumour Type and Control Notation	Result Type	Avg No of Tissue Samples	SDev of Sample No	Fold Diff from normal	SDev of Fold or % diff.	No of Replicate Experiments		
PAG1	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0	0.0	3.0	0.0	2.0		
PAG1	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	0.0	4.7	0.6	3.0		
PAG1	U133A & B	Melanoma All Vs Essential Normal	80-50 Percentile Ratio	41.0	0.0	6.0	1.7	3.0		
PAG1	U133A & B	Melanoma All Vs Essential Normal	80-80 Percentile Ratio	41.0	0.0	3.5	0.7	2.0		

Table 38

Gene Symbol	Array Type	Tumour Type and Control Notation	Result Type	Avg No of Tissue Samples	% samples showing diff level of expression	No of Replicate Experiments
PDGFRA	U133A	Breast Cancer	<=0.5X Concordance	50.0	60.0	1.0

	& B						
PDGFRA	U133A & B	Colon Cancer		<=0.5X Concordance	27.0	30.0	1.0
PDGFRA	U133A & B	Colon Cancer		2X Concordance	27.0	30.0	1.0
PDGFRA	U133A & B	Colon metastasis		<=0.5X Concordance	33.0	42.0	1.0
PDGFRA	U133A & B	Colon metastasis		2X Concordance	32.0	38.0	1.0

Table 39

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
PLAU	Chiron cDNA array	Breast Cancer	2X Concordance	21.8	2.5	38.0	2.0	4.0
PLAU	Chiron cDNA array	Colon Cancer	2X Concordance	58.3	19.0	34.8	11.6	6.0
PLAU	Chiron	Colon metastasis	2X Concordance	31.8	1.6	28.6	5.4	5.0

	cDNA array									
PLAU	Chiron cDNA array	matched Colon Cancer	2X Concordance	36.0	0.0	36.0	15	4.0		
PLAU	Chiron cDNA array	matched Colon Cancer metastasis	2X Concordance	30.2	10.7	37.4	17.0	5.0		
PLAU	Chiron cDNA array	Prostate Cancer	2X Concordance	106.0		41.0		1.0		
PLAU	AffyEVD Array Type	Colon Cancer	2X Concordance	27.0		37.0		1.0	No of Replicate Experiments	
PLAU	U133A & B	Breast Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	43.0	0.0	2.0	0.0	2.0	SDev of Fold or % diff.	
PLAU	U133A & B	Lung Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	68.0	0.0	4.5	0.7	2.0		
PLAU	U133A & B	Lung Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	68.0	0.0	4.0	0.0	2.0		
PLAU	U133A	Pancreas Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	13.0	0.0	5.5	0.7	2.0		

	& B								
PLAU	U133A & B	Pancreas Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	13.0	0.0	4.5	0.7	2.0	
PLAU	U133A & B	Upper Aero-Digestive Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	23.0	0.0	4.0	0.0	2.0	
PLAU	U133A & B	Upper Aero-Digestive Tract Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	23.0	0.0	5.0	0.0	2.0	

Table 40

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
PTPRK	U133A & B	Breast Cancer	<=0.5X Concordance	50.0	24.0	1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>Fold Diff from normal</u>	<u>No of Replicate Experiments</u>
PTPRK	U133A & B	Pancreas Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	13.0	3.0	1.0
PTPRK	U133A & B	Renal Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	66.0	3.0	1.0
PTPRK	U133A & B	Stomach Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	23.0	3.0	1.0

PTPRK	U133A & B	Upper Aero-Digestive Tract Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	23.0	3.0	1.0
PTPRK	U133A & B	Uterus Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	39.0	3.0	1.0

Table 41

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
RAMP1	U133A & B	Colon metastasis	2X Concordance	33.0	21.0	1.0
RAMP1	U133A & B	Prostate Cancer	2X Concordance	22.0	27.0	1.0
RAMP1	U133A & B	Prostate Cancer vs adjacent Normal	2X Concordance	16.0	31.0	1.0

Table 42

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
RSPO3	Chiron cDNA array	Breast Cancer	<=0.5X Concordance	23.0	22.0	1.0

Table 43

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
SELL	Chiron cDNA array	Colon Cancer	2X Concordance	77.0	35.0	1.0
SELL	Chiron cDNA array	Colon metastasis	2X Concordance	33.0	27.0	1.0
SELL	Chiron cDNA array	matched Colon Cancer	2X Concordance	36.0	28.0	1.0
SELL	Chiron cDNA array	matched Colon Cancer metastasis	2X Concordance	36.0	42.0	1.0
SELL	U133A & B	Prostate Cancer	2X Concordance	9.0	100.0	1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>Fold Diff. from normal</u>	<u>No of Replicate Experiments</u>
SELL	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0	27.0	1.0
SELL	U133A	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	38.0	1.0

	& B					
--	-----	--	--	--	--	--

Table 44

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
SLAMF6	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0	5.0	1.0
SLAMF6	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	7.0	1.0

Table 45

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
SLC12A7	Chiron cDNA array	Breast Cancer	2X Concordance	20.5	3.5	24.0	2.8	2.0
SLC12A7	Chiron cDNA array	Colon Cancer	2X Concordance	19.0		21.0		1.0
SLC12A7	Chiron	matched Colon Cancer	2X Concordance	18.0		22.0		1.0

	cDNA array								
SLC12A7	Chiron cDNA array	matched Colon Cancer metastasis	2X Concordance	18.0			28.0		1.0
SLC12A7	U133A & B	Breast Cancer	2X Concordance	50.0			22.0		1.0
Gene Symbol	Array Type	Tumour Type and Control Notation	Result Type	Avg No of Tissue Samples	SDev of Sample No	% samples showing diff level of expression	SDev of Fold or % diff.	No of Replicate Experiments	
SLC12A7	U133A & B	Liver Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	21.0		3.0		1.0	
SLC12A7	U133A & B	Renal Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	66.0		3.0		1.0	

Table 46

Gene Symbol	Array Type	Tumour Type and Control Notation	Result Type	Avg No of Tissue Samples	% samples showing diff level of expression	No of Replicate Experiments
SLC1A2	U133A & B	Colon Cancer	<=0.5X Concordance	26.0	54.0	1.0

Table 47

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
SLC9A9	U133A & B	Colon Cancer	<=0.5X Concordance	11.0		64.0		1.0
SLC9A9	U133A & B	Prostate Cancer	<=0.5X Concordance	17.0		29.0		1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>Fold Diff from normal</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
SLC9A9	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	0.0	2.5	0.7	2.0

Table 48

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
SLPI	Chiron cDNA array	Breast Cancer	<=0.5X Concordance	23.0	0.0	41.5	9.2	2.0
SLPI	Chiron cDNA array	Breast Cancer	2X Concordance	23.0		22.0		1.0
SLPI	Chiron	Colon Cancer	<=0.5X	70.5	9.2	24.5	3.5	2.0

	cDNA array		Concordance						
SLPI	Chiron cDNA array	Colon Cancer	2X Concordance	77.0				35.0	1.0
SLPI	Chiron cDNA array	Colon metastasis	$\leq 0.5X$ Concordance	33.0				42.0	1.0
SLPI	Chiron cDNA array	Colon metastasis	2X Concordance	33.0				36.0	1.0
SLPI	Chiron cDNA array	matched Colon Cancer	2X Concordance	30.0	8.5			34.5	2.0
SLPI	Chiron cDNA array	matched Colon Cancer Cancer	$\leq 0.5X$ Concordance	36.0				22.0	1.0
SLPI	Chiron cDNA array	matched Colon Cancer metastasis	$\leq 0.5X$ Concordance	27.3	7.5			30.7	3.0
SLPI	Chiron cDNA array	matched Colon Cancer metastasis	2X Concordance	36.0				25.0	1.0

	array									
SLPI	Chiron cDNA array	Prostate Cancer	<=0.5X Concordance	106.0				25.0		1.0
SLPI	U133A & B	Breast Cancer	<=0.5X Concordance	50.0				54.0		1.0
SLPI	U133A & B	Colon Cancer	<=0.5X Concordance	21.0				43.0		1.0
SLPI	U133A & B	Colon metastasis	<=0.5X Concordance	33.0				45.0		1.0
Gene Symbol	Array Type	Tumour Type and Control Notation	Result Type	Avg No of Tissue Samples	SDev of Sample No	Fold Diff from normal	SDev of Fold or % diff.	No of Replicate Experiments		
SLPI	U133A & B	Lung Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	68.0		26.0		1.0		
SLPI	U133A & B	Lung Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	68.0		14.0		1.0		
SLPI	U133A & B	Ovarian Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	72.0		91.0		1.0		
SLPI	U133A & B	Ovarian Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	72.0		49.0		1.0		

SLPI	U133A & B	Uterus Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	39.0		131.0		1.0
SLPI	U133A & B	Uterus Endometastasisrial Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	39.0		71.0		1.0

Table 49

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
TGFBR3	Chiron cDNA array	Breast Cancer	<=0.5X Concordance	23.0	57.0	1.0
TGFBR3	Chiron cDNA array	matched Colon Cancer	2X Concordance	36.0	28.0	1.0
TGFBR3	U133A & B	Breast Cancer	<=0.5X Concordance	50.0	72.0	1.0
TGFBR3	U133A & B	Colon Cancer	2X Concordance	24.0	25.0	1.0
TGFBR3	U133A & B	Prostate Cancer	<=0.5X Concordance	22.0	23.0	1.0
TGFBR3	U133A	Prostate Cancer vs adjacent Normal	<=0.5X Concordance	16.0	31.0	1.0

	& B						
--	-----	--	--	--	--	--	--

Table 50

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
THY1	U133A & B	Colon Cancer	2X Concordance	25.0		28.0		1.0
THY1	U133A & B	Colon metastasis	2X Concordance	33.0		27.0		1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>Fold Diff from normal</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
THY1	U133A & B	Pancreas Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	13.0	0.0	4.7	1.2	3.0
THY1	U133A & B	Pancreas Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	13.0		3.0		1.0

Table 51

Gene Symbol	Array Type	Tumour Type and Control Notation	Result Type	Avg No of Tissue Samples	Fold Diff from normal	No of Replicate Experiments
-------------	------------	----------------------------------	-------------	--------------------------	-----------------------	-----------------------------

TNFRSF14	U133A & B	Melanoma All Vs Essential Normal	80-50 Percentile Ratio	41.0	2.0	1.0
----------	-----------	----------------------------------	------------------------	------	-----	-----

Table 52

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>Fold Diff from normal</u>	<u>No of Replicate Experiments</u>
TNFSF11	U133A & B	Blood Lymph Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	60.0	5.0	1.0
TNFSF11	U133A & B	Blood/Lymph Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	60.0	7.0	1.0
TNFSF11	U133A & B	Pancreas Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	13.0	4.0	1.0

Table 53

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
TPCN1	U133A & B	Colon metastasis	<=0.5X Concordance	33.0	27.0	1.0
TPCN1	U133A & B	Prostate Cancer vs adjacent Normal	<=0.5X Concordance	16.0	25.0	1.0

Table 54

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
TSPAN2	U133A & B	Colon Cancer	<=0.5X Concordance	26.0	27.0	1.0
TSPAN2	U133A & B	Prostate Cancer vs adjacent Normal	2X Concordance	15.0	20.0	1.0

Table 55

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
WNT1	U133A & B	Prostate Cancer	2X Concordance	21.0	24.0	1.0

Table 56

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
WNT3	U133A & B	Colon Cancer	<=0.5X Concordance	26.0	58.0	1.0

Table 57

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>% samples showing diff level of expression</u>	<u>No of Replicate Experiments</u>
WNT3A	Chiron cDNA array	Breast Cancer	2X Concordance	23.0	22.0	1.0

Table 58

<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>% samples showing diff level of expression</u>	<u>Fold Diff from normal</u>	<u>SDev of Fold or diff.</u>
PRLR	U133A&B	Breast Cancer	<=0.5X Concordance	49.0		20.0		1.0
PRLR	U133A&B	Colon Cancer	<=0.5X Concordance	26.5	0.7	33.5	14.8	2.0
PRLR	U133A&B	Colon metastasis	<=0.5X Concordance	31.0		29.0		1.0
<u>Gene Symbol</u>	<u>Array Type</u>	<u>Tumour Type and Control Notation</u>	<u>Result Type</u>	<u>Avg No of Tissue Samples</u>	<u>SDev of Sample No</u>	<u>Fold Diff from normal</u>	<u>SDev of Fold or % diff.</u>	<u>No of Replicate Experiments</u>
PRLR	U133+2	Breast Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	43.0	0.0	6.3	3.0	4.0
PRLR	U133+2	Breast Carcinoma All Vs Essential Normal	80-80 Percentile Ratio	43.0	0.0	4.0	1.0	3.0
PRLR	U133+2	Liver Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	21.0	0.0	3.5	0.7	2.0
PRLR	U133+2	Stomach Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	23.0		4.0		1.0
PRLR	U133+2	Uterus Carcinoma All Vs Essential Normal	80-50 Percentile Ratio	39.0		4.0		1.0

[illegible]

Example 14: Functional Assays

[00493] SiRNAs targeting listed genes were manufactured on 96-well plates. Three different functional assays, proliferation assay, cytotoxicity assay, and soft agar growth assay, were performed after reverse transfection in different cell lines. 1-3 replicates were generated per assay per cell line. Primers and probes used for the assays are set forth in Table 2, above.

Statistical analysis procedure

[00494] Cytotoxic effect: To determine significance of cytotoxic effect, proliferation assay readouts were ranked in ascending order for each replicate. Soft agar assay readouts were ranked in ascending order for each replicate. Cytotoxicity assay readouts were ranked in descending order for each replicate. The rank for the same target was summated among all the replicates of all the assays. Labels of targets were randomly permuted 100,000 times and the rank-sum of each individual target was recorded. The experimental rank-sum was compared versus the distribution of the randomly permuted rank-sum of each target to obtain p-values of “cytotoxic” effect.

[00495] Cytotoxic effect: To determine significance of cytotoxic effect, proliferation assay readouts were ranked in ascending order for each replicate. Soft agar assay readouts were ranked in ascending order for each replicate. Cytotoxicity assay readouts were ranked in descending order for each replicate. The rank for the same target was summated among all the replicates of all the assays. Labels of targets were randomly permuted 100,000 times and the rank-sum of each individual target was recorded. The experimental rank-sum was compared versus the distribution of the randomly permuted rank-sum of each target to obtain p-values of “cytostatic” effect

[00496] Out of control effect: To determine significance of out of control effect, proliferation assay readouts were ranked in ascending order for each replicate. Soft agar assay readouts were ranked in ascending order for each replicate. Cytotoxicity assay readouts were ranked in descending order for each replicate. The rank for the same target was summated among all the replicates of all the assays. Labels of targets were randomly permuted 100,000 times and the rank-sum of each individual target was recorded. The experimental rank-sum was compared versus the distribution of the randomly permuted rank-sum of each target to obtain p-values of “out of control” effect.

[00497] Ineffective effect: To determine significance of cytotoxic effect, proliferation assay readouts were ranked in ascending order for each replicate. Soft agar assay readouts were

ranked in ascending order for each replicate. Cytotoxicity assay readouts were ranked in descending order for each replicate. The rank for the same target was summated among all the replicates of all the assays. Labels of targets were randomly permuted 100,000 times and the rank-sum of each individual target was recorded. The experimental rank-sum was compared versus the distribution of the randomly permuted rank-sum of each target to obtain p-values of “ineffective” effect.

[00498] Significant calls: Minimum p-value among four effects for each target were determined. If the minimum p-value < 0.05, a corresponding significant call was assigned. If the minimum p-value \geq 0.05, a non-significant call was assigned.

[00499] Results are set forth in Table 59, below

Table 59

Individual	PC3	A549	HT29
ATP1A2	ineffective	ineffective	ineffective
B2M	cytotoxic	ineffective	ineffective
CD38	ineffective	non-significant	ineffective
CPM	ineffective	non-significant	ineffective
DPT	cytostatic	cytotoxic	cytotoxic
FGF10	out of control	cytotoxic	non-significant
FGF4	cytotoxic	cytotoxic	cytotoxic
FGF8	non-significant		non-significant
IGSF8	cytostatic		ineffective
IL4R	cytotoxic	cytostatic	non-significant
IL6R	non-significant	non-significant	ineffective
ITGA4	out of control	non-significant	out of control
LFNG	non-significant	cytotoxic	cytostatic
LTB	ineffective	out of control	out of control
NOTCH1	non-significant	out of control	non-significant
NOTCH2	out of control	non-significant	non-significant
PAG1	cytostatic		non-significant
PLAU	non-significant	non-significant	non-significant
PRG1	cytotoxic		cytostatic
PTPRK	non-significant	out of control	ineffective
SLC12A7	non-significant		non-significant
SLPI	out of control		non-significant
THY1	out of control		cytotoxic

Legend: Cytotoxic: P value significant for anti-proliferative and cytotoxic effect
 Cytostatic: P value significant for cytostatic but not cytotoxic effect
 out of control: cytotoxicity significant but anti-proliferative effect weak or absent
 insignificant: p value for cytostatic or cytotoxic effect NOT significant
 ineffective: no anti-proliferative or cytotoxic effect

[00500] All publications, patents and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

[00501] It will be understood that the invention has been described by way of example only and modifications may be made whilst remaining within the scope and spirit of the invention.

CLAIMS

1. A method for treating cancer in a patient comprising modulating the level of an expression product of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFB3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, wherein the cancer is selected from the group consisting of melanoma, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, bladder cancer, stomach cancer or skin cancer.
2. The method of claim 1 wherein said method comprises administering to the patient an antibody, a nucleic acid, or a polypeptide that modulates the level of said gene expression product.
3. The method according to claim 1 or 2 wherein the expression level of the expression product is upregulated or downregulated by at least a 2-fold change.
4. The method according to any one of claims 1 to 3 wherein the cancer is treated by the inhibition of tumour growth or the reduction of tumour volume.
5. The method according to any one of claims 1 to 4 wherein the cancer is treated by reducing the invasiveness of a cancer cell.
6. The method of claim 1 wherein the expression product is a protein or mRNA.
7. The method according to claim 6, wherein the level of the expression product at a first time point is compared to the level of the same expression product at a second time point, wherein an increase in level of the expression product at the second time point relative to the first time point is indicative of the progression of cancer.
8. The method according to claim 2 wherein the nucleotide has a sequence selected from the group consisting of SEQ ID NOS:1-144.
9. The method of claim 2 wherein the antibody is a neutralizing antibody.

10. The method of claim 2 wherein the antibody is a monoclonal antibody.
11. The method of claim 2 wherein the antibody is a monoclonal antibody which binds to a polypeptide encoded by said gene with an affinity of at least $1 \times 10^8 K_a$.
12. The method of claim 2 wherein the antibody is a monoclonal antibody, a polyclonal antibody, a chimeric antibody, a human antibody, a humanized antibody, a single-chain antibody, a bi-specific antibody, a multi-specific antibody, or a Fab fragment.
13. A method of treating a cancer in a patient characterized by overexpression of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, relative to a control, the method comprising modulating expression of the gene in the patient.
14. The method of claim 13 wherein said method comprises administering to the patient an antibody, a nucleic acid, or a polypeptide that inhibits gene expression.
15. A method for diagnosing cancer comprising detecting evidence of differential expression in a patient sample of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, wherein evidence of differential expression is diagnostic of cancer, wherein the cancer is selected from the group consisting of melanoma, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, bladder cancer, stomach cancer or skin cancer.

16. The method of claim 15 wherein evidence of differential expression is detected by measuring the level of an expression product of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAUI, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA.
17. The method of claim 16 wherein the expression product is a protein or mRNA.
18. The method of claim 17 wherein the level of expression of protein is measured using an antibody which specifically binds to a polypeptide encoded by the gene.
19. The method of claim 18 wherein the antibody is linked to an imaging agent.
20. The method of claim 16 wherein the level of expression product of the gene in the patient sample is compared to a control.
21. The method of claim 24 wherein the control is a known normal tissue of the same tissue type as in the patient sample.
22. The method of claim 20 wherein the level of the expression product in the sample is increased relative to the control.
23. A method for detecting a cancerous cell in a patient sample comprising detecting evidence of an expression product of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAUI, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, wherein evidence of expression of the gene in the sample indicates that a cell in the sample is cancerous.

24. The method of claim 23 wherein the cell is a breast cell, colon cell, kidney cell, liver cell, lung cell, lymphatic cell, ovary cell, pancreas cell, prostate cell, uterine cell, cervical cell, bladder cell, stomach cell or skin cell.
25. The method of claim 23 wherein evidence of the expression product is detected using an antibody linked to an imaging agent.
26. A method for assessing the progression of cancer in a patient comprising comparing the level of an expression product of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAUI, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, in a biological sample at a first time point to a level of the same expression product at a second time point, wherein a change in the level of the expression product at the second time point relative to the first time point is indicative of the progression of the cancer, wherein the cancer is selected from the group consisting of melanoma, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, bladder cancer, stomach cancer or skin cancer.
27. A method of diagnosing cancer selected from the group consisting of melanoma, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, bladder cancer, stomach cancer or skin cancer, the method comprising:

(a) measuring a level of mRNA of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAUI, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and

PDGFRA, in a first sample, said first sample comprising a first tissue type of a first individual; and

(b) comparing the level of the mRNA in (a) to:

(1) a level of the mRNA in a second sample, said second sample comprising a normal tissue type of said first individual, or

(2) a level of the mRNA in a third sample, said third sample comprising a normal tissue type from an unaffected individual;

wherein at least a two fold difference between the level of mRNA in (a) and the level of the mRNA in the second sample or the third sample indicates that the first individual has or is predisposed to cancer.

28. The method of claim 27 wherein at least a three fold difference between the level of mRNA in (a) and the level of the mRNA in the second sample or the third sample indicates that the first individual has or is predisposed to cancer.

29. A method of screening for anti-cancer activity comprising:

(a) contacting a cell that expresses a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAUI, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, with a candidate anti-cancer agent; and

(b) detecting at least a two fold difference between the level of gene expression in the cell in the presence and in the absence of the candidate anti-cancer agent,

wherein at least a two fold difference between the level of gene expression in the cell in the presence and in the absence of the candidate anti-cancer agent indicates that the candidate anti-cancer agent has anti-cancer activity, wherein the cancer is selected from the group consisting of melanoma, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, bladder cancer, stomach cancer or skin cancer.

30. The method of claim 29 wherein at least a three fold difference between the level of gene expression in the cell in the presence and in the absence of the candidate anti-cancer agent indicates that the candidate anti-cancer agent has anti-cancer activity.
31. The method of claim 29 wherein the candidate anti-cancer agent is an antibody, small organic compound, small inorganic compound, or polynucleotide.
32. The method of claim 31 wherein the polynucleotide is an antisense oligonucleotide.
33. A method for identifying a patient as susceptible to treatment with an antibody that binds to an expression product of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA, comprising measuring the level of the expression product of the gene in a biological sample from that patient.
34. A kit for the diagnosis or detection of cancer in a mammal, wherein said kit comprises an antibody or fragment thereof, or an immunoconjugate or fragment thereof, according to any one of the proceeding embodiments, wherein said antibody or fragment specifically binds a tumor cell antigen of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA; one or more reagents for detecting a binding reaction between said antibody and said tumor cell antigen; and optionally instructions for using the kit, wherein the cancer is selected from the group consisting of melanoma, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer,

ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, bladder cancer, stomach cancer or skin cancer.

35. A kit for diagnosing cancer comprising a nucleic acid probe that hybridises under stringent conditions to a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA; primers for amplifying the gene; and optionally instructions for using the kit, wherein the cancer is selected from the group consisting of melanoma, breast cancer, colon cancer, kidney cancer, liver cancer, lung cancer, ovary cancer, pancreatic cancer, prostate cancer, uterine cancer, cervical cancer, bladder cancer, stomach cancer or skin cancer.

36. A composition comprising one or more antibodies or oligonucleotides specific for an expression product of a gene selected from the group consisting of WNT3, NOTCH1, WNT1, WNT3A, IL2RA, FLT3, PRLR, TNFRSF6, PTPRK, TNFSF11, FGFR3, SELL, DPT, CD28, PDCD1LG2, LFNG, ANXA2, CDH23, IFNAR2, RAMP1, MET, FGFR1, PLAU, SLC4A8, TPCN1, LTBR, IL4R, ITGA9, FGFR2, FCMD, COL15A1, SLAMF6, CD74, TSPAN-2, THSD2, CD2, LNPEP, IGSF8, ATP1A2, ITGA4, B2M, PAG, NOTCH2, UTS2, TNFRSF14, TGFBR3, FGF8, FGF3, XYLT1, FGF4, LOC390243, THY1, SLC9A9, PRG1, PRLR, CPM, FBN2, LTB, KCNH5, SLC12A7, FGF10, IL7R, MATN2, COL19A1, SLC1A2, SLC28A2, SLPI, IL12A, IL6R, TNFSF8, TNFRSF8, CD38 and PDGFRA.

37. The composition of claim 36 further comprising a conventional cancer medicament.

38. The composition of claim 36 further comprising a pharmaceutically acceptable excipient.

39. The composition of claim 36 wherein the one or more oligonucleotides are selected from the group consisting of SEQ ID NOS:1-144.